

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic nucleic search, using sw model
Run on: August 9, 2001, 10:25:47; Search time: 62.54 seconds
(without alignments)
4752.343 Million cell updates/sec

Title: US-09-475-704-4
Perfect score: 1509
Sequence: 1 ataggccacgacacacat.....cagacccacgacacacat 1509

Scoring table: IDENTITY_NBC
Gapop 10.0, Gapext 1.0

Searched: 730101 seqs, 41950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2690909090

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:	Nucleotide	DB ID	Length	DB ID	Description
1:	/cml_9/qqqdat/a/geneseq/geneseq/NA1980.DAT	1509	21	AAA51610	HIV synthetic Gag
2:	/cml_9/qqqdat/a/geneseq/geneseq/NA1981.DAT	1509	21	AAA51626	HIV codon optimized
3:	/cml_9/qqqdat/a/geneseq/geneseq/NA1982.DAT	1509	21	AAA51609	HIV synthetic Gag
4:	/cml_9/qqqdat/a/geneseq/geneseq/NA1983.DAT	1509	21	AAA51625	HIV codon optimized
5:	/cml_9/qqqdat/a/geneseq/geneseq/NA1984.DAT	1509	21	AAA51625	Synthetic HIV Gag
6:	/cml_9/qqqdat/a/geneseq/geneseq/NA1985.DAT	1509	21	AAA51625	HIV biotronic co
7:	/cml_9/qqqdat/a/geneseq/geneseq/NA1986.DAT	1509	21	AAA51625	HIV biotronic co
8:	/cml_9/qqqdat/a/geneseq/geneseq/NA1987.DAT	1509	21	AAA51625	HIV biotronic co
9:	/cml_9/qqqdat/a/geneseq/geneseq/NA1988.DAT	1509	21	AAA51625	Synthetic HIV Gag
10:	/cml_9/qqqdat/a/geneseq/geneseq/NA1989.DAT	1509	21	AAA51625	HIV Gag-protease c
11:	/cml_9/qqqdat/a/geneseq/geneseq/NA1990.DAT	1509	21	AAA51625	
12:	/cml_9/qqqdat/a/geneseq/geneseq/NA1991.DAT	1509	21	AAA51625	
13:	/cml_9/qqqdat/a/geneseq/geneseq/NA1992.DAT	1509	21	AAA51625	
14:	/cml_9/qqqdat/a/geneseq/geneseq/NA1993.DAT	1509	21	AAA51625	
15:	/cml_9/qqqdat/a/geneseq/geneseq/NA1994.DAT	1509	21	AAA51625	
16:	/cml_9/qqqdat/a/geneseq/geneseq/NA1995.DAT	1509	21	AAA51625	
17:	/cml_9/qqqdat/a/geneseq/geneseq/NA1996.DAT	1509	21	AAA51625	
18:	/cml_9/qqqdat/a/geneseq/geneseq/NA1997.DAT	1509	21	AAA51625	
19:	/cml_9/qqqdat/a/geneseq/geneseq/NA1998.DAT	1509	21	AAA51625	
20:	/cml_9/qqqdat/a/geneseq/geneseq/NA1999.DAT	1509	21	AAA51625	
21:	/cml_9/qqqdat/a/geneseq/geneseq/NA2000.DAT	1509	21	AAA51625	
22:	/cml_9/qqqdat/a/geneseq/geneseq/NA2001.DAT	1509	21	AAA51625	

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	1509	100.0	1509	21	AAA51610
2	1489.8	98.7	1509	21	AAA51626
3	1476.8	84.6	1479	21	AAA51609
4	1450.8	84.6	1479	21	AAA51625
5	1201	79.6	1472	21	AAA50412
6	1201	79.6	1472	21	AAA50472
7	1201	79.6	1472	21	AAA50473
8	1201	79.6	1472	21	AAA50471
9	1201	79.6	1472	21	AAA50470
10	1194	79.5	201	21	AAA70415
11	1117.8	74.1	1853	21	AAA70413

SUMMARIES

12	1117.8	74.1	1865	21	AAA70475	HIV Gag-protease c
13	1117.8	74.1	1865	21	AAA70476	HIV Gag-protease c
14	1117.8	74.1	4319	21	AAA70414	HIV Gag-polymerase
15	1077.4	71.4	1503	21	AA252051	Codon optimised HIV
16	1077.4	71.4	8908	21	AA252055	Packaging construct
17	1049.2	69.5	1268	21	AAA70417	HIV Gag common seq
18	971.4	64.4	1532	22	AA312843	Humanised HIV-1 gag
19	969.8	64.3	1532	19	AAV45374	Synthetic HIV gag
20	930.2	61.6	4307	20	AAZ08740	HIV gagpol SYN-pc
21	930.2	61.6	4307	21	AAV49722	Human immunodeficient
22	930.2	61.6	4307	22	AAV48786	Nucleotide sequence
23	925.4	61.3	4327	21	AAA94984	HIV partial leader
24	925.4	61.3	4327	21	AAA94984	HIV partial leader
25	925.4	61.3	4327	21	AAA94984	HIV complete leader
26	916.6	60.7	1482	22	AAI41286	Humanised HIV-1 gag
27	734.2	48.7	8972	21	AAI57927	HIV-1 non-subtype
28	709.6	47.0	9009	21	AAI57929	HIV-1 non-subtype
29	709.6	47.0	9009	21	AAI57926	HIV-1 non-subtype
30	663.6	44.0	8992	21	AAI57922	HIV-1 non-subtype
31	654	43.3	9143	11	AAQ06635	Complete sequence
32	651.0	43.2	8968	21	AAI57919	HIV-1 non-subtype
33	646	42.8	9193	20	AAV04986	HIV-1-JC infection
34	642.2	42.6	1503	19	AAV35386	HIV-1 gag protein
35	642.2	42.6	1503	21	AA252050	Wild type human im
36	642.2	42.6	7899	14	AAQ34479	HPACK1. Human im
37	642.2	42.6	7899	20	AAV04704	HPACK1 DNA const
38	642.2	42.6	9709	13	AAQ22488	HIV-1 proviral clo
39	642.2	42.6	9709	16	AAQ06140	HIV-1 NL4-3 genom
40	642.2	42.6	9709	20	AAV01871	Nucleic acid seque
41	642.2	42.6	9709	21	AAV0926	HIV-1 env DNA. Hu
42	642.2	42.6	9709	21	AAV0928	HIV-1 viral protei
43	642.2	42.6	9709	21	AAV0928	HIV-1 RNA-1118 rec
44	642.2	42.6	12494	20	AAV04206	Vector pHP-1 compl
45	642.2	42.6	15561	18	AAV14455	Plasmid pNAISGLP

ALIGNMENTS

RESULT 1	
AAA51610	
10	AAA51610 standard: DNA: 1509 BP.
XX	
AC	AAA51610:
XX	
XX	31-SEP-2000 (first entry)
XX	
XX	HIV synthetic Gag polynucleotide.
XX	
XX	Gag, expression cassette, antigenic, type C, HIV, env, 8/ethet loc
XX	DNA immunization, packaging cell line, antigen presentation; SS.
XX	
XX	Human immunodeficiency virus type C strain AF110947.
XX	Synthetic.
XX	
XX	W320009304-A2.
XX	
XX	06-JUL-2000.
XX	
XX	30-DEC-1999; 99NO-US31273.
XX	
XX	31-DEC-1998; 98US-0114495.
XX	01-SEP-1999; 99US-0152195.
XX	(CHIR) CHIRON CORP.
XX	
XX	Barnett S., Zur Megede J.
XX	WPI: 2000-452401/39.
XX	
XX	Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
XX	env polypeptide and the polypeptide useful for immunizing a mammal
XX	especially human against HIV

[illegible]

PS Claim 1; Page 92; 113pp; English.
 XX Expression cassettes comprising a polynucleotide encoding antigenic
 CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
 CC useful in DNA immunization, generation of packaging cell lines and
 CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
 CC expression cassettes exhibit increased potency for induction of
 CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV 1
 CC self-assemble into non-infectious virus-like particles which are used as
 CC a matrix for the proper presentation of an antigen entrapped or
 CC associated to the immune system of the host.
 XX Sequence 60 BP; 14 A; 22 C; 15 G; 9 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcaagcaggccccaagagcccttcgcgactacgtgagcgtttcttcaagacc 60
 DB 1 gacatcaagcaggccccaagagcccttcgcgactacgtgagcgtttcttcaagacc 60

RESULT 2
 AAA51609
 ID AAA51609 standard; DNA; 1479 BP.
 XX
 AC AAA51609;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE HIV synthetic Gag polynucleotide.
 XX
 KW Gag; expression cassette; antigenic; type C, HIV. Env; synthetic;
 KW DNA immunization; packaging cell line; antigen presentation; ss.
 XX
 OS Human immunodeficiency virus type C strain AF110965.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1479
 FT /*tag= a
 FT /product= Synthetic_Gag
 FT /note= "Codon usage pattern was modified and inhibitory
 FT elements (INS) and RRE sites were inactivated
 FT resulting in improved expression"
 XX

PN WO200039304-A2.
 XX
 XX 06-JUL-2000.
 XX
 XX 30-DEC-1999; 99WO-US31273.
 XX
 XX 31-DEC-1998; 98US-0114495.
 XX 01-SEP-1999; 99US-0152195.
 XX
 XX (CHTR) CHIRON CORP.
 XX
 XX Barnett S, Zur Megede J;
 XX
 XX WPI; 2000-452401/39.
 XX
 XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
 XX Env polypeptide and the polypeptide useful for immunizing a mammal
 XX especially human against HIV
 XX

PS Claim 2; Page 92-93; 113pp; English.
 XX Expression cassettes comprising a polynucleotide encoding antigenic
 CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
 CC useful in DNA immunization, generation of packaging cell lines and

CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
 CC expression cassettes exhibit increased potency for induction of
 CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
 CC self-assemble into non-infectious virus-like particles which are used as
 CC a matrix for the proper presentation of an antigen entrapped or
 CC associated to the immune system of the host.
 XX
 XX Sequence 1479 BP; 325 A; 529 C; 463 G; 162 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 1479;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcaagcaggccccaagagcccttcgcgactacgtgagcgtttcttcaagacc 60
 DB 814 gacatcaagcaggccccaagagcccttcgcgactacgtgagcgtttcttcaagacc 903

RESULT 3
 AAA51625
 ID AAA51625 standard; DNA; 1479 BP.
 XX
 AC AAA51625;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE HIV codon-optimized synthetic Gag polynucleotide.
 XX
 KW Gag; expression cassette; antigenic; type C, HIV. Env; synthetic;
 KW DNA immunization; packaging cell line; antigen presentation; ss.
 XX
 OS Human immunodeficiency virus type C strain AF110965.
 OS Synthetic.
 XX
 XX WO200039304-A2.
 XX
 XX 06-JUL-2000.
 XX
 XX 30-DEC-1999; 99WO-US31273.
 XX
 XX 31-DEC-1998; 98US-0114495.
 XX 01-SEP-1999; 99US-0152195.
 XX
 XX (CHTR) CHIRON CORP.
 XX
 XX Barnett S, Zur Megede J;
 XX
 XX WPI; 2000-452401/39.

XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
 XX Env polypeptide and the polypeptide useful for immunizing a mammal
 XX especially human against HIV
 XX
 XX Disclosure; Page 103-104; 113pp; English.
 XX
 XX Expression cassettes comprising a polynucleotide encoding antigenic
 XX type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
 XX useful in DNA immunization, generation of packaging cell lines and
 XX production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
 XX expression cassettes exhibit increased potency for induction of
 XX cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
 XX self-assemble into non-infectious virus-like particles which are used as
 XX a matrix for the proper presentation of an antigen entrapped or
 XX associated to the immune system of the host.
 XX
 XX Sequence 1479 BP; 325 A; 533 C; 461 G; 160 T; 0 other;

Query Match 100.0%; Score 60; DB 21; Length 1479;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

XX (CHIR ) CHIRON CORP.
XX PA
XX XX
XX PPI
XX XX
XX WPI: 2000-452401/39
XX DR
XX PT
XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV
XX PPT Env polypeptide and the polypeptide useful for immunizing a mammal
XX PPT especially human against HIV
XX XX
XX Disclosure: Page 104; 113pp; English.
XX XX
XX Expression cassettes comprising a polynucleotide encoding antigenic
XX CCC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are
XX CCC useful in DNA immunization, generation or packaging cell lines and
XX CCC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag
XX CCC expression cassettes exhibit increased potency for induction of
XX CCC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1
XX CCC self-assemble into non-infectious virus-like particles which are used as
XX CCC a matrix for the proper presentation of an antigen entrapped or
XX CCC associated to the immune system of the host.
XX Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 other;
XX SQ
Query Match: 92.0%; Score 55.2; LB 21; Length 1509;
Best Local Similarity 95.0%; Prod. No. 6-9e-09;
Matches 57; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 qacatacgaagcaggcccccagacaccccttcgcgactacgtgagccgtttcttcaagacc 60
Db 1111111111111111111111111111111111111111111111111111111111111111
641 qacatccgcagcaggcccccagagcccttcgcgactacgtgagccgtttcttcaagacc 960
RESULT 7
AAA70426
ID AAA70426 standard; DNA; 60 BP.
XX AC
XX AAA70426;
XX 28-NOV-2000 (first entry)
XX DT
XX XX
XX HIV p55 Gag Major Homology Region coding sequence.
DE
XX XX
XX HIV-1; AIDS; Gag; vaccine; expression cassette; ss.
XX KW
XX Human immunodeficiency virus type 1.
XX OS
XX FH
XX Key Location/Qualifiers
XX CDS 1..60
XX FT FT /*tag= a
XX FT FT /partial
XX FT FT /product= "HIV p55 Gag Major Homology Region protein;"
XX FT FT /note= "No stop codon given"
XX PP
XX WO200039302-A2.
XX XX
XX 06-JUL-2000.
XX XX
XX 30-DEC-1999; 99WO-US31245.
XX PF
XX 31-DEC-1998; 98US-0114495.
XX PF 01-DEC-1999; 99US-0168471.
XX PR
XX XX
XX (CHIR ) CHIRON CORP.
XX PA
XX Barnett S., Zur Megede J., Srivastava L., Lian Y., Hartog K., Liu H.;
XX PI Greer C., Selby M., Walker C.;
XX XX
XX WPI: 2000-452400/39.
XX DR P-PSDB; AAB14214.
XX XX

```

CC infection and acquired immunodeficiency syndrome (AIDS). The present
 CC sequence is a common region found in Gag coding sequences. Gag protease
 CC and Gag-polymerase
 XX
 CC Sequence 1268 BP; 274 A; 449 C; 411 G; 135 T; 0 other;
 SQ

Query Match 89.3%; Score 53.6; DB 21; Length 1268;
 Best Local Similarity 94.3%; Pred. No. 2.2e-08;
 Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 qacatcaagcaagcccaaggaagccttcacgaactacatgaacccctcttcaagaac 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 862 qacatccagcaagcccaaggaagccttcacgaactacatgaacccctcttcaagaac 921

RESULT 9
 AAZ52051
 ID AAZ52051 standard; DNA; 1503 BP.
 AC AAZ52051;
 XX
 DT 18 JUL 2000 (first entry)
 XX
 DE Codon optimised human immunodeficiency virus gag coding region.
 XX
 KW HIV; gag; packaging cell line; lentivirus; retroviral vector particle;
 KW gag; gene therapy; gene replacement; vaccine; biochemical reagent;
 KW codon optimisation; ds.
 XX
 KW Human immunodeficiency virus.
 OS
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 CHS 1..1503
 F1 /*tag a
 F1 /product= "gag protein"
 XX
 W0200015819-A1.
 XX
 XX 24-MAR-2000.
 XX
 PF 10 SEP-1999; 99WO-US0675.
 XX
 PR 11 SEP-1998; 98US-010022.
 PR 12 SEP-1998; 98US-010063.
 XX
 XX (CHIL) CHILDEPENS MEDICAL CENT.
 PA
 XX Gray JL, Mulligan RC;
 F1
 XX WPI; 2000-273455/23.
 DB
 PR PSQC AAZ70544
 XX

New packaging cell line for producing a viral accessory protein
 F1 independent HIV derived retroviral vector particles, useful in gene
 F1 therapy or gene replacement -
 XX
 PS Claim 41; Fig 8; 62pp; English.
 XX

The patent discloses new packaging cell line for producing a viral
 CC accessory protein independent lentiviruses, preferably
 CC human immunodeficiency virus (HIV), derived retroviral vector particles.
 CC The packaging cell line comprises a mammalian cell, a retroviral DNA
 CC comprising a coding sequence for a lentivirus, preferably HIV, gagpol,
 CC where the coding sequence has been mutated to improve expression of
 CC the viral gagpol proteins, a second retroviral nucleotide sequence
 CC comprising the coding sequence for a heterologous envelope protein and a
 CC third retroviral nucleotide sequence comprising a DNA sequence of
 CC interest and lentiviruses, preferably HIV, cis-acting sequences required
 CC for packaging, reverse transcription and integration.
 CC The packaging cell lines and viral particles can be used for gene
 CC therapy or gene replacement with improved safety. They can also be used

CC in the development and production of vaccines and biochemical reagents.
 CC the present sequence is a gag coding region of codon optimised HIV
 CC gagpol sequence. This sequence is used in the packaging
 CC cell line. Codon optimisation results in improved expression of the
 CC gagpol protein and reduces the risk of recombination between the transfer
 CC vector and gagpol mRNA.
 XX
 SQ Sequence 1503 BP; 339 A; 530 C; 433 G; 201 T; 0 other;

Query Match 89.3%; Score 53.6; DB 21; Length 1503;
 Best Local Similarity 94.3%; Pred. No. 2.2e-08;
 Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 qacatcaagcaagcccaaggaagccttcacgaactacatgaacccctcttcaagaac 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 850 qacatccagcaagcccaaggaagccttcacgaactacatgaacccctcttcaagaac 909

RESULT 10
 AAZ70412
 ID AAZ70412 standard; DNA; 1515 BP.
 AC AAZ70412;
 XX
 DT 28-NOV-2000 (first entry)
 XX
 DE Synthetic HIV Gag expression cassette coding sequence Gag.ModSP2.
 DE HIV-1; AIDS; Gag; vaccine; expression cassette; ss.
 KW
 KW Human immunodeficiency virus type 1.
 OS
 OS Synthetic.
 XX
 PN W0200049302-A2.
 XX
 PD 06-JUL-2000.
 XX
 PF 30-DEC-1999; 99WO-US01245.
 XX
 PR 31-DEC-1998; 98US-0114495.
 PR 01-DEC-1999; 99US-0168471.
 XX
 XX (CHIL) CHIRON CORP.
 XX
 PI Barnett S, Zar Medode J, Srivastava L, Lian Y, Bartoa K, Liu H;
 PI Greer C, Selby M, Walker C;
 XX
 DR WPI; 2500-452406/49.
 XX
 PT Expression cassettes encoding the human immunodeficiency virus (HIV)
 PT Gag containing polypeptide useful for vaccinating against HIV
 PT infections and acquired immunodeficiency syndrome (AIDS) -
 XX
 PS Claim 3; Fig 7; 49pp; English.
 XX

The present sequence is the coding sequence of a HIV Gag expression
 CC cassette, Gag.ModSP2. The Gag protein of HIV is needed for the assembly
 CC of virus-like particles. In addition, the gag protein is involved in
 CC many stages of the HIV life cycle, including assembly, virion maturation
 CC after particle release and early post-entry steps in viral replication.
 CC The expression cassette may be used for the recombinant expression of
 CC HIV Gag-polypeptides which may then be used to vaccinate against HIV
 CC infection and acquired immunodeficiency syndrome (AIDS).
 XX
 SQ Sequence 1515 BP; 329 A; 547 C; 480 G; 159 T; 0 other;

Query Match 89.3%; Score 53.6; DB 21; Length 1515;
 Best Local Similarity 94.3%; Pred. No. 2.2e-08;
 Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 qacatcaagcaagcccaaggaagccttcacgaactacatgaacccctcttcaagaac 60


```

XX Barnett S, Zur Mequde J, Srivastava I, Lian Y, Hartog K, Liu H;
PI Greer C, Selby M, Walker C;
XX WPI: 2000-452400/49.
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PI Gag containing polypeptide useful for vaccinating against HIV
PI infections and acquired immunodeficiency syndrome (AIDS) .
XX Claim 5; Fig 70; 391pp; English.
XX The present sequence is the coding sequence of a HIV gag-protease
CC expression cassette, GaPol.ModSF. The Gag protein of HIV is
CC needed for the assembly of virus-like particles. In addition, the Gag
CC protein is involved in many stages of the HIV life cycle, including
CC assembly, virion maturation and early post-entry steps in viral
CC steps in viral replication. The expression cassette may be used for the
CC recombinant expression of HIV Gag-polypeptides which may then be used to
CC vaccinate against HIV infection and acquired immunodeficiency syndrome
CC (AIDS).
XX Sequence 1865 BP; 460 A; 583 C; 569 G; 253 T; 0 other;
SQ
Query Match 89.4%, Score 53.6; DB 21; Length 1865;
Best Local Similarity 93.3%; Pred. No. 2.4e-08;
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 1 qacatcagacagagcccaagagagccttcgcgactacatqgaccccttctcaagacc 60
PI [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]
DB 868 qacatcagacagagcccaagagagccttcgcgactacatqgaccccttctcaagacc 927

RESULT 14
AAA70415
ID AAA70415 standard; DNA; 2041 BP.
XX AAA70415;
XX 28-NOV-2000 (first entry)
XX Synthetic HIV Gag/HCV core fusion coding sequence.
XX HIV-1; AIDS; Gag; vaccine; expression cassette; ss.
XX ChimERIC - Human immunodeficiency virus type 1.
XX ChimERIC - Hepatitis C virus.
XX W0200039302-A2.
XX 06-JUL-2000.
XX 30-DEC-1999; 99WO-US31245.
XX 31-DEC-1998; 98US-0114495.
XX 01-DEC-1999; 99US-0168471.
XX (CHIR ) CHIRON CORP.
XX Barnett S, Zur Mequde J, Srivastava I, Lian Y, Hartog K, Liu H,
PI Greer C, Selby M, Walker C;
XX WPI: 2000-452400/49.
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PI Gag containing polypeptide useful for vaccinating against HIV
PI infections and acquired immunodeficiency syndrome (AIDS) .
XX Claim 5; Fig 70; 391pp; English.
XX The present sequence is a HIV Gag/hepatitis C virus (HCV) core fusion
CC coding sequence. The Gag protein of HIV is needed for the assembly of

```

```

CC virus-like particles. In addition, the Gag protein is involved in many
CC stages of the HIV life cycle, including assembly, virion maturation after
CC particle release and early post-entry steps in viral replication. The
CC present invention relates to synthetic HIV Gag expression cassettes. The
CC present sequence was cloned and used to generate the expression cassettes
CC of the present invention. The expression cassettes may be used for the
CC recombinant expression of HIV Gag-polypeptides which may then be used to
CC vaccinate against HIV infection and acquired immunodeficiency syndrome
CC (AIDS).
XX Sequence 2031 BP; 421 A; 707 C; 646 G; 257 T; 0 other;
SQ
Query Match 89.4%, Score 53.6; DB 21; Length 2031;
Best Local Similarity 93.3%; Pred. No. 2.4e-08;
Matches 56; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 1 qacatcagacagagcccaagagagccttcgcgactacatqgaccccttctcaagacc 60
PI [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]
DB 862 qacatcagacagagcccaagagagccttcgcgactacatqgaccccttctcaagacc 921

RESULT 15
AAA70414
ID AAA70414 standard; DNA; 4319 BP.
XX AAA70414;
XX 28-NOV-2000 (first entry)
XX HIV Gag-polymerase expression cassette coding sequence GaPol.ModSF.
XX HIV-1; AIDS; Gag polymerase; vaccine; expression cassette; ss.
XX Human immunodeficiency virus type 1.
XX Synthetic.
XX W0200039302-A2.
XX 06-JUL-2000.
XX 30-DEC-1999; 99WO-US31245.
XX 31-DEC-1998; 98US-0114495.
XX 01-DEC-1999; 99US-0168471.
XX (CHIR ) CHIRON CORP.
XX Barnett S, Zur Mequde J, Srivastava I, Lian Y, Hartog K, Liu H;
PI Greer C, Selby M, Walker C;
XX WPI: 2000-452400/49.
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PI Gag containing polypeptide useful for vaccinating against HIV
PI infections and acquired immunodeficiency syndrome (AIDS) .
XX Claim 10; Fig 7; 391pp; English.
XX The present sequence is the coding sequence of a HIV Gag-polymerase
CC expression cassette, GaPol.ModSF. The Gag protein of HIV is needed for
CC the assembly of virus-like particles. In addition, the Gag protein is
CC involved in many stages of the HIV life cycle, including assembly, virion
CC maturation after particle release and early post-entry steps in viral
CC replication. The expression cassette may be used for the recombinant
CC expression of HIV Gag polypeptides which may then be used to vaccinate
CC against HIV infection and acquired immunodeficiency syndrome (AIDS).
XX Sequence 4319 BP; 1003 A; 1426 C; 1361 G; 529 T; 0 other;
SQ
Query Match 89.4%; Score 53.6; DB 21; Length 4319;
Best Local Similarity 93.3%; Pred. No. 2.5e-08;

```

	Matches	56;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
Qy	1	gacatcgaagcaagggcccccagagagcccttcgcgactacgtggagccgcttcttcaagacc	60							
Db	862	gacatcgcgcagggcccccagagagcccttcgcgactacgtggagccgcttcttcaagacc	921							

Search completed. August 8, 2001, 18:24:58
Job time: 12602 sec

Query Match
Best Local Similarity 71.7%; Score 43; DB 9; Length 4307;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 qacatcaaaaggaaggccgaaagaacattcttcgcqgaatacgttggagacgtttcttcadgaac 59
||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 850 GACATCGGCAAGGGCCGAAGGAAAGCTTTTCGTGACTAGTGGAGTGGTCTTATAAAAA 908

RESULT 10

AX035465	LOCUS	AX035465	4327 bp	DNA	PAT	15 NOV-2000
DEFINITION	Sequence 2 from Patent WO005541.					
ACCESSION	AX035465					
VERSION	AX035465.1	GI:11191095				
KEYWORDS	Synthetic construct, artificial sequence,					
SOURCE	synthetic construct					
ORGANISM	artificial sequence,					
REFERENCE	1 (bases 1 to 4307)					
AUTHORS	Oden,M. and Mitrophanous,K.					
TITLE	Anti-viral vectors					
JOURNAL	Patent: WO 005541-A 2 21-SEP 2000;					
FEATURES	UDEN MARK (GB) : OXFORD BIOMEDICA LTD (GB) ; MITROPHANOUS KYRIACOS (US) Location/Qualifiers 1..4307 /organism:"synthetic construct" /db_xref:"taxon:42640" /note-"SynGP4 - codon optimised gagpol sequence"					
BASE COUNT	1135 a 1198 c 1285 g 689 t					
ORIGIN						
Query Match	71.7%; Score 43; DB 9; Length 4307;					
Best Local Similarity	83.1%; Pred. No. 0.0066;					
Matches	49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;					

QY 1 qacatcaaaaggaaggccgaaagaacattcttcgcqgaatacgttggagacgtttcttcadgaac 59
||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 850 GACATCGGCAAGGGCCGAAGGAAAGCTTTTCGTGACTAGTGGAGTGGTCTTATAAAAA 908

RESULT 9

AX056846	LOCUS	AX056846	4307 bp	DNA	PAT	17-JAN-2001
DEFINITION	Sequence 2 from Patent W00075370.					
ACCESSION	AX056846					
VERSION	AX056846.1	GI:12309774				
KEYWORDS	Synthetic construct, synthetic construct					
SOURCE	synthetic construct					
ORGANISM	artificial sequence,					
REFERENCE	1 (bases 1 to 4307)					
AUTHORS	Mitrophanous,K., Kim,N.H. and Katsouridou,E.					
TITLE	In vitro selection method for determining inhibitory rna molecules					
JOURNAL	Patent: W0 007537-A 2 11 JAN 2001;					
FEATURES	UDEN MARK (GB) : OXFORD BIOMEDICA LTD (GB) Location/Qualifiers 1..4307 /organism:"synthetic construct" /db_xref:"taxon:42640" /note-"codon optimised gagpol sequence"					
BASE COUNT	1135 a 1198 c 1285 g 689 t					
ORIGIN						
Query Match	71.7%; Score 43; DB 9; Length 4307;					
Best Local Similarity	83.1%; pred. No. 0.0066;					
Matches	49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;					

```

RESULT 12
AX035463
LOCUS       AX035463             4642 bp    DNA
DEFINITION   Sequence 12 from Patent WO0005541.
ACCESSION   AX035463
VERSION     AX035463.1  GI:11191105
KEYWORDS    synthetic construct.
            artificial construct.
REFERENCE    1 (bases 1 to 4642)
AUTHORS     Uden.M. and Mitrophanous.K.
TITLE       Anti-viral vectors
JOURNAL     Eur J Cell Biol 163:1-11 (1999)
FEATURES    Location/Qualifiers
             source
               1..4642
               /organism="synthetic construct"
               /db_xref="taxon:32630"
               /note="PSYNG82 - codon optimised HIV-1 gagpol with leader
               sequence"
BASE COUNT   1219 a 1273 c 1389 g 762 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

/product="p24"
/protein_id="CAB85850.1"
/db_xref="GI:7414162"
/translation="MNIQQQVYQATSPRTLNAMVKVIEKAFSPVIMPSALAEARA
TQDINTMLNTVCGHCAAMQILKDTINEFAAEHDLHPVHAGVPAGQIREPRGSDIA
GTSTLQEOITWMTSNPPIVGEIYKGIILGLAKIYPMYSPVSIIDIKGKPEPRD
YVDRFKTLRAEQSSQEVKSMWMTDLLIONANPOCKTLLKALGASGASLEMMTACQV
GRP"
BASE COUNT   243 a 128 c 154 g 139 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

Query Match      65.3%; Score 39.2; DB 59; Length 663;
Best Local Similarity 78.3%; Pred. No. 0.13;
Matches 47; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

BASE COUNT   145 a 70 c 97 g 81 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
9
```

117:	qb_est48.*
118:	qb_est49.*
119:	qb_est50.*
120:	qb_est51.*
121:	qb_est52.*
122:	qb_est53.*
123:	qb_est54.*
124:	qb_est55.*
125:	qb_est56.*
126:	qb_est57.*
127:	qb_est58.*
128:	qb_est59.*
129:	qb_est60.*
130:	qb_est61.*
131:	qb_est62.*
132:	qb_est63.*
133:	qb_est64.*
134:	qb_est65.*
135:	qb_est66.*
136:	qb_est67.*
137:	qb_est68.*
138:	qb_est69.*
139:	qb_est70.*
140:	qb_est71.*
141:	qb_est72.*
142:	qb_est73.*
143:	qb_est74.*
144:	qb_est75.*
145:	qb_est76.*
146:	qb_est77.*
147:	qb_est78.*
148:	qb_est79.*
149:	qb_est80.*
150:	qb_est81.*
151:	qb_est82.*
152:	qb_est83.*
153:	qb_est84.*
154:	qb_est85.*
155:	qb_est86.*
156:	qb_est87.*
157:	qb_est88.*
158:	qb_est89.*
159:	qb_est90.*
160:	qb_est91.*
161:	qb_est92.*
162:	qb_est93.*
163:	qb_est94.*
164:	qb_est95.*
165:	qb_est96.*
166:	qb_est97.*
167:	qb_est98.*
168:	qb_est99.*
169:	qb_est100.*
170:	qb_est101.*
171:	qb_est102.*
172:	qb_est103.*
173:	qb_est104.*
174:	qb_est105.*
175:	qb_est106.*
176:	qb_est107.*
177:	qb_est108.*
178:	qb_est109.*
179:	qb_est110.*
180:	qb_est111.*
181:	qb_est112.*
182:	qb_est113.*
183:	qb_est114.*
184:	qb_est115.*
185:	qb_est116.*
186:	qb_est117.*
187:	qb_est118.*
188:	qb_est119.*
189:	qb_est120.*
190:	qb_est121.*
191:	qb_est122.*
192:	qb_est123.*
193:	qb_est124.*
194:	qb_est125.*
195:	qb_est126.*
196:	qb_est127.*
197:	qb_est128.*
198:	qb_est129.*
199:	qb_est130.*
200:	qb_est131.*
201:	qb_est132.*
202:	qb_est133.*
203:	qb_est134.*
204:	qb_est135.*
205:	qb_est136.*
206:	qb_est137.*
207:	qb_est138.*
208:	qb_est139.*
209:	qb_est140.*
210:	qb_est141.*
211:	qb_est142.*
212:	qb_est143.*
213:	qb_est144.*
214:	qb_est145.*
215:	qb_est146.*
216:	qb_est147.*
217:	qb_est148.*
218:	qb_est149.*
219:	qb_est150.*
220:	qb_est151.*
221:	qb_est152.*
222:	qb_est153.*
223:	qb_est154.*
224:	qb_est155.*
225:	qb_est156.*
226:	qb_est157.*
227:	qb_est158.*
228:	qb_est159.*
229:	qb_est160.*
230:	qb_est161.*
231:	qb_est162.*
232:	qb_est163.*
233:	qb_est164.*
234:	qb_est165.*
235:	qb_est166.*
236:	qb_est167.*
237:	qb_est168.*
238:	qb_est169.*
239:	qb_est170.*
240:	qb_est171.*
241:	qb_est172.*
242:	qb_est173.*
243:	qb_est174.*
244:	qb_est175.*
245:	qb_est176.*
246:	qb_est177.*
247:	qb_est178.*
248:	qb_est179.*
249:	qb_est180.*
250:	qb_est181.*
251:	qb_est182.*
252:	qb_est183.*
253:	qb_est184.*
254:	qb_est185.*
255:	qb_est186.*
256:	qb_est187.*
257:	qb_est188.*
258:	qb_est189.*
259:	qb_est190.*
260:	qb_est191.*
261:	qb_est192.*
262:	qb_est193.*
263:	qb_est194.*
264:	qb_est195.*
265:	qb_est196.*
266:	qb_est197.*
267:	qb_est198.*
268:	qb_est199.*
269:	qb_est200.*
270:	qb_est201.*
271:	qb_est202.*
272:	qb_est203.*
273:	qb_est204.*
274:	qb_est205.*
275:	qb_est206.*
276:	qb_est207.*
277:	qb_est208.*
278:	qb_est209.*
279:	qb_est210.*
280:	qb_est211.*
281:	qb_est212.*
282:	qb_est213.*
283:	qb_est214.*
284:	qb_est215.*
285:	qb_est216.*
286:	qb_est217.*
287:	qb_est218.*
288:	qb_est219.*
289:	qb_est220.*
290:	qb_est221.*
291:	qb_est222.*
292:	qb_est223.*
293:	qb_est224.*
294:	qb_est225.*
295:	qb_est226.*
296:	qb_est227.*
29	

190:	gb_est110:*
191:	gb_est111:*
192:	gb_hic:*
193:	em_gss_fun1:*
194:	em_gss_fun2:*
195:	em_gss_hum3:*
196:	em_gss_hum3:*
197:	em_gss_hum4:*
198:	em_gss_hum5:*
199:	em_gss_hum6:*
200:	em_gss_hum7:*
201:	em_gss_hum8:*
202:	em_gss_hum9:*
203:	em_gss_inv1:*
204:	em_gss_inv2:*
205:	em_gss_inv3:*
206:	em_gss_inv3:*
207:	em_gss_pln1:*
208:	em_gss_pln2:*
209:	em_gss_pro:*
210:	em_gss_r0d1:*
211:	em_gss_r0d2:*
212:	em_gss_r0d3:*
213:	em_gss_r0d4:*
214:	em_gss_r0d5:*
215:	em_gss_vrt1:*
216:	em_gss_vrt2:*
217:	em_gss_vrt3:*
218:	qb_gss1:*
219:	qb_gss2:*
220:	qb_gss3:*
221:	qb_gss4:*
222:	qb_gss5:*
223:	qb_gss6:*
224:	qb_gss7:*
225:	qb_gss8:*
226:	qb_gss9:*
227:	qb_gss10:*
228:	qb_gss11:*
229:	qb_gss12:*
230:	qb_gss13:*
231:	qb_gss14:*
232:	qb_gss15:*
233:	qb_gss16:*
234:	qb_gss17:*
235:	qb_gss18:*
236:	qb_gss19:*
237:	qb_gss20:*
238:	qb_gss21:*
239:	qb_gss22:*
240:	qb_gss23:*
241:	qb_gss24:*
242:	qb_gss25:*
243:	qb_gss26:*
244:	qb_gss27:*
245:	qb_gss28:*
246:	qb_gss29:*
247:	qb_gss30:*
248:	qb_gss31:*
249:	qb_gss32:*
250:	qb_gss33:*
251:	qb_gss34:*
252:	em_gss_inv4:*
253:	em_gss_r0d6:*
254:	em_gss_r0d7:*
255:	qb_gss35:*
256:	qb_gss36:*
257:	qb_gss37:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	28.6	47.7	491	167	AB011268	AB011268
2	28.4	47.3	560	154	BC517417	BC517417
3	28.2	47.0	585	112	AW147097	AW147097
4	28	46.7	168	174	BC050418	BC050418
5	28	46.7	216	2	AA143927	AA143927
6	28	46.7	219	166	BE356040	BE356040
7	28	46.7	527	147	BE417794	BE417794
8	28	46.7	536	23	AI586607	AI586607
9	28	46.7	447	167	AB011268	AB011268
10	28	46.7	455	167	AB011268	AB011268
11	28	46.7	463	156	C59276	C59276
12	28	46.7	464	141	BE918440	BE918440
13	28	46.7	464	188	T14732	T14732
14	28	46.7	470	137	BE595775	BE595775
15	28	46.7	471	141	BE917843	BE917843
16	28	46.7	472	154	BC462293	BC462293
17	28	46.7	487	174	BC047940	BC047940
18	28	46.7	489	151	BF657016	BF657016
19	28	46.7	491	174	BC048475	BC048475
20	28	46.7	497	175	BC273239	BC273239
21	28	46.7	498	166	BE364444	BE364444
22	28	46.7	498	175	BE249984	BE249984
23	28	46.7	497	151	BF657119	BF657119
24	28	46.7	498	175	BC273239	BC273239
25	28	46.7	498	175	BC273239	BC273239
26	28	46.7	498	175	BC273239	BC273239
27	28	46.7	498	175	BC273239	BC273239
28	28	46.7	498	175	BC273239	BC273239
29	28	46.7	498	175	BC273239	BC273239
30	28	46.7	498	175	BC273239	BC273239
31	28	46.7	498	175	BC273239	BC273239
32	28	46.7	498	175	BC273239	BC273239
33	28	46.7	498	175	BC273239	BC273239
34	28	46.7	498	175	BC273239	BC273239
35	28	46.7	498	175	BC273239	BC273239
36	28	46.7	498	175	BC273239	BC273239
37	28	46.7	498	175	BC273239	BC273239
38	28	46.7	498	175	BC273239	BC273239
39	28	46.7	498	175	BC273239	BC273239
40	28	46.7	498	175	BC273239	BC273239
41	28	46.7	498	175	BC273239	BC273239
42	28	46.7	498	175	BC273239	BC273239
43	28	46.7	498	175	BC273239	BC273239
44	28	46.7	498	175	BC273239	BC273239
45	28	46.7	498	175	BC273239	BC273239

ALIGNMENTS

Result No.	Score	Query Match	Length	DB	ID	Description
1	28.6	47.7	491	167	AB011268	AB011268
2	28.4	47.3	560	154	BC517417	BC517417
3	28.2	47.0	585	112	AW147097	AW147097
4	28	46.7	168	174	BC050418	BC050418
5	28	46.7	216	2	AA143927	AA143927
6	28	46.7	219	166	BE356040	BE356040
7	28	46.7	527	147	BE417794	BE417794
8	28	46.7	536	23	AI586607	AI586607
9	28	46.7	447	167	AB011268	AB011268
10	28	46.7	455	167	AB011268	AB011268
11	28	46.7	463	156	C59276	C59276
12	28	46.7	464	141	BE918440	BE918440
13	28	46.7	464	188	T14732	T14732
14	28	46.7	470	137	BE595775	BE595775
15	28	46.7	471	141	BE917843	BE917843
16	28	46.7	472	154	BC462293	BC462293
17	28	46.7	487	174	BC047940	BC047940
18	28	46.7	489	151	BF657016	BF657016
19	28	46.7	491	174	BC048475	BC048475
20	28	46.7	497	175	BC273239	BC273239
21	28	46.7	498	166	BE364444	BE364444
22	28	46.7	498	175	BE249984	BE249984
23	28	46.7	497	151	BF657119	BF657119
24	28	46.7	498	175	BC273239	BC273239
25	28	46.7	498	175	BC273239	BC273239
26	28	46.7	498	175	BC273239	BC273239
27	28	46.7	498	175	BC273239	BC273239
28	28	46.7	498	175	BC273239	BC273239
29	28	46.7	498	175	BC273239	BC273239
30	28	46.7	498	175	BC273239	BC273239
31	28	46.7	498	175	BC273239	BC273239
32	28	46.7	498	175	BC273239	BC273239
33	28	46.7	498	175	BC273239	BC273239
34	28	46.7	498	175	BC273239	BC273239
35	28	46.7	498	175	BC273239	BC273239
36	28	46.7	498	175	BC273239	BC273239
37	28	46.7	498	175	BC273239	BC273239
38	28	46.7	498	175	BC273239	BC273239
39	28	46.7	498	175	BC273239	BC273239
40	28	46.7	498	175	BC273239	BC273239
41	28	46.7	498	175	BC273239	BC273239
42	28	46.7	498	175	BC273239	BC273239
43	28	46.7	498	175	BC273239	BC273239
44	28	46.7	498	175	BC273239	BC273239
45	28	46.7	498	175	BC273239	BC273239

COMMENT

Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program
2-1-2 Kannondai, Tsukuba
Ibaraki,
Japan 305

Tel.: 0298-48-7441
Fax: 0298-48-7468

Email: tsasaki@affrc.affrc.go.jp

PROJECT: "RGP"

Location/Qualifiers

1. 491

Organism: "oryza sativa"

Strain: "Nipponbare"

Lib: "X80" "X80" "X80"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

COMMENT

Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program
2-1-2 Kannondai, Tsukuba
Ibaraki,
Japan 305

Tel.: 0298-48-7441
Fax: 0298-48-7468

Email: tsasaki@affrc.affrc.go.jp

PROJECT: "RGP"

Location/Qualifiers

1. 491

Organism: "oryza sativa"

Strain: "Nipponbare"

Lib: "X80" "X80" "X80"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"

Accession: "E1244_22"


```

KEYWORDS      EST
SOURCE        Zea mays.
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
              clade; Panicoidae; Andropogoneae; Zea.
REFERENCE     1 (bases 1 to 336)
AUTHORS      Walbot V.
TITLE        Maize ESTs from various cDNA libraries sequenced at Stanford
              University
              Unpublished (1999)
              Contact: Walbot V
              Department of Biological Sciences
              Stanford University
              855 California Ave, Palo Alto, CA 94304, USA
              Tel: 650 723 3227
              Fax: 650 725 8221
              Email: walbot@stanford.edu
              Plate: 486049 row: D column: 10.
FEATURES      Location/Qualifiers
              1..336
              /organism="Zea mays"
              /cultivar="B73"
              /db_xref="taxon:4577"
              /clone_lib="486 - leaf primordia cDNA library from Hake
              lab"
              /tissue_type="leaf primordia"
              /dev_stage="p7-p11 leaf"
              /lab_host="E.coli XL1-Blue MFR"
              /notes="organ: shoot; Vector: Lambda zap; Hake lab cDNA
              library."
BASE COUNT    82 a 91 c 101 g 62 t
ORIGIN
Query Match  46.7%; Score 28; DB 22; Length 336;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 40; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 gacatcaagcaggccccaaggagcccttcgcgactacgtgagccgtcttctcaagacc 60
DB 313 GTCAATCAAGAGAGACCCGACATCAAGAGAGCTTCACGAGCTCTCTCTCTCAAGAAC 254

RESULT 9
LOCUS      AU097510 347 bp mRNA EST 30-JUN-2000
DEFINITION AU097510 Rice shoot Oryza sativa cDNA clone S5505, mRNA sequence.
ACCESSION  AU097510
VERSION     AU097510.1 GI:8860192
KEYWORDS    EST.
SOURCE      Oryza sativa.
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzoaceae; Oryza.
              1 (bases 1 to 347)
              Rice cDNA from etiolated shoot (2000)
              Unpublished (2000)
              Contact: Takuji Sasaki
              National Institute of Agrobiological Resources
              Rice Genome Research Program
              2-1-2 Kannondai, Tsukuba
              Ibaraki,
              Japan 305
              Tel: 0298-38-7441
              Fax: 0298-38-7468
              Email: tsasaki@abr.affrc.go.jp
              PROJECT="RGP"
              S5505_87.
FEATURES      Location/Qualifiers
              1..347
              /organism="Oryza sativa"
              /strain="Nipponbare"
              /db_xref="taxon:4530"
              /clone_lib="E31104_6Z"
              /dev_stage="shorter than 3cm"
              /note="organ: panicle"
BASE COUNT    94 a 97 c 96 g 67 t 1 others
ORIGIN
Query Match  46.7%; Score 28; DB 107; Length 355;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 40; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 gacatcaagcaggccccaaggagcccttcgcgactacgtgagccgtcttctcaagacc 60
DB 69 GTCAATCAAGAGAGACCCGACATCAAGAGAGCTTCACGAGCTCTCTCTCTCAAGAAC 128

RESULT 11
LOCUS      C99276 463 bp mRNA EST 19-OCT-1998
DEFINITION C99276 Rice panicle at ripening stage Oryza sativa cDNA clone
              E10581_6Z, mRNA sequence.
FEATURES      Location/Qualifiers
              1..463
              /organism="Oryza sativa"
              /strain="Nipponbare"
              /db_xref="taxon:4530"
              /clone_lib="E31104_6Z"
              /dev_stage="shorter than 3cm"
              /note="organ: panicle"
BASE COUNT    96 a 97 c 96 g 67 t 1 others
ORIGIN

```

```

/organism="Oryza sativa"
/strain="Nipponbare, sub_species Japonica"
/db_xref="taxon:4530"
/clone="S5505"
/clone_lib="Rice shoot"
/note="Etiolated shoot (8 days old)"
BASE COUNT    86 a 95 c 95 g 70 t 1 others
ORIGIN
Query Match  46.7%; Score 28; DB 107; Length 347;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 40; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 gacatcaagcaggccccaaggagcccttcgcgactacgtgagccgtcttctcaagacc 60
DB 63 GTCAATCAAGAGAGACCCGACATCAAGAGAGCTTCACGAGCTCTCTCTCTCAAGAAC 122

RESULT 10
LOCUS      AU029576 355 bp mRNA EST 19-OCT-1998
DEFINITION AU029576 Rice panicle shorter than 3cm Oryza sativa cDNA clone
              E31104_6Z, mRNA sequence.
ACCESSION  AU029576
VERSION     AU029576.1 GI:3762824
KEYWORDS    EST.
SOURCE      Oryza sativa.
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzoaceae; Oryza.
              1 (bases 1 to 355)
              Sasaki, T. and Yamamoto, K.
              Rice cDNA from panicle
              Unpublished (1997)
              Contact: Takuji Sasaki
              National Institute of Agrobiological Resources
              Rice Genome Research Program
              2-1-2 Kannondai, Tsukuba
              Ibaraki,
              Japan 305
              Tel: 0298-38-7441
              Fax: 0298-38-7468
              Email: tsasaki@abr.affrc.go.jp
              PROJECT="RGP"
              E10581_6Z.
FEATURES      Location/Qualifiers
              1..355
              /organism="Oryza sativa"
              /strain="Nipponbare"
              /db_xref="taxon:4530"
              /clone_lib="E31104_6Z"
              /dev_stage="shorter than 3cm"
              /note="organ: panicle"
BASE COUNT    94 a 97 c 96 g 67 t 1 others
ORIGIN
Query Match  46.7%; Score 28; DB 107; Length 355;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 40; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 gacatcaagcaggccccaaggagcccttcgcgactacgtgagccgtcttctcaagacc 60
DB 69 GTCAATCAAGAGAGACCCGACATCAAGAGAGCTTCACGAGCTCTCTCTCTCAAGAAC 128

RESULT 11
LOCUS      C99276 463 bp mRNA EST 19-OCT-1998
DEFINITION C99276 Rice panicle at ripening stage Oryza sativa cDNA clone
              E10581_6Z, mRNA sequence.
FEATURES      Location/Qualifiers
              1..463
              /organism="Oryza sativa"
              /strain="Nipponbare"
              /db_xref="taxon:4530"
              /clone_lib="E31104_6Z"
              /dev_stage="shorter than 3cm"
              /note="organ: panicle"
BASE COUNT    94 a 97 c 96 g 67 t 1 others
ORIGIN

```


Matches	40;	Conservative	0;	Mismatches	20;	Indels	0;	Gaps	0;
57	1	qatfcacatcatatqccacaaqaaqaapcccttccacgaactacatqqaacccctctctccaaqacc	60						
106	81	GTGATTAAGGATAGAGGATAGCAAGAGAGTCTACGATCTCGAGGACTATGCTGCTCAAGAAC	140						

Search completed: August 8, 2001, 20:58:57
Job time: 21841 sec

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000


```

117: qb_est48:*
118: qb_est49:*
119: qb_est50:*
120: qb_est51:*
121: qb_est52:*
122: qb_est53:*
123: qb_est54:*
124: qb_est55:*
125: qb_est56:*
126: qb_est57:*
127: qb_est58:*
128: qb_est59:*
129: qb_est60:*
130: qb_est61:*
131: qb_est62:*
132: qb_est63:*
133: qb_est64:*
134: qb_est65:*
135: qb_est66:*
136: qb_est75:*
137: qb_est76:*
138: qb_est77:*
139: qb_est78:*
140: qb_est79:*
141: qb_est80:*
142: qb_est81:*
143: qb_est82:*
144: qb_est83:*
145: qb_est84:*
146: qb_est85:*
147: qb_est86:*
148: qb_est87:*
149: qb_est88:*
150: qb_est89:*
151: qb_est90:*
152: qb_est99:*
153: qb_est100:*
154: qb_est101:*
155: qb_est102:*
156: qb_est103:*
157: qb_est104:*
158: qb_est105:*
159: qb_est106:*
160: qb_est167:*
161: qb_est168:*
162: qb_est169:*
163: qb_est170:*
164: qb_est171:*
165: qb_est172:*
166: qb_est173:*
167: qb_est174:*
168: qb_est191:*
169: qb_est192:*
170: qb_est193:*
171: qb_est194:*
172: qb_est195:*
173: qb_est196:*
174: qb_est197:*
175: qb_est198:*
176: qb_est199:*
177: qb_est200:*
178: qb_est201:*
179: qb_est202:*
180: qb_est203:*
181: qb_est204:*
182: qb_est205:*
183: qb_est206:*
184: qb_est207:*
185: qb_est208:*
186: qb_est209:*
187: qb_est210:*
188: qb_est211:*
189: qb_est212:*

```

```

190: qb_est110:*
191: qb_est111:*
192: qb_est112:*
193: qb_est113:*
194: qb_est114:*
195: qb_est115:*
196: qb_est116:*
197: qb_est117:*
198: qb_est118:*
199: qb_est119:*
200: qb_est120:*
201: qb_est121:*
202: qb_est122:*
203: qb_est123:*
204: qb_est124:*
205: qb_est125:*
206: qb_est126:*
207: qb_est127:*
208: qb_est128:*
209: qb_est129:*
210: qb_est130:*
211: qb_est131:*
212: qb_est132:*
213: qb_est133:*
214: qb_est134:*
215: qb_est135:*
216: qb_est136:*
217: qb_est137:*
218: qb_est138:*
219: qb_est139:*
220: qb_est140:*
221: qb_est141:*
222: qb_est142:*
223: qb_est143:*
224: qb_est144:*
225: qb_est145:*
226: qb_est146:*
227: qb_est147:*
228: qb_est148:*
229: qb_est149:*
230: qb_est150:*
231: qb_est151:*
232: qb_est152:*
233: qb_est153:*
234: qb_est154:*
235: qb_est155:*
236: qb_est156:*
237: qb_est157:*
238: qb_est158:*
239: qb_est159:*
240: qb_est160:*
241: qb_est161:*
242: qb_est162:*
243: qb_est163:*
244: qb_est164:*
245: qb_est165:*
246: qb_est166:*
247: qb_est167:*
248: qb_est168:*
249: qb_est169:*
250: qb_est170:*
251: qb_est171:*
252: qb_est172:*
253: qb_est173:*
254: qb_est174:*
255: qb_est175:*
256: qb_est176:*
257: qb_est177:*
258: qb_est178:*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



PS Claim 1: Page 92; 113pp; English

XX Expression cassettes comprising a polynucleotide encoding antigenic

CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are

CC useful in DNA immunization, generation of packaging cell lines and

CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag

CC expression cassettes exhibit increased potency for induction of

CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1

CC self-assemble into non-infectious virus-like particles which are used as

CC a matrix for the proper presentation of an antigen entrapped or

CC associated to the immune system of the host.

XX Sequence 60 BP; 12 A; 24 C; 15 G; 9 T; 0 other;

SQ

Query Match 100.0%; Score 60; DB 21; Length 60;

Best Local Similarity 100.0%; Pred. No. 7 2e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatccgagggcccaaggagcccttcgcgactacgtgagccgtttcttcaagacc 60

DB 1 gacatccgagggcccaaggagcccttcgcgactacgtgagccgtttcttcaagacc 60

RESULT 2

AAA51610

ID AAA51610 standard; DNA; 1509 BP.

AC AAA51610;

XX 31-OCT-2000 (first entry)

XX HIV synthetic Gag polynucleotide.

DE

XX Gag; expression cassette; antigenic, type C; HIV; Env; synthetic;

KW DNA immunization; packaging cell line; antigen presentation; ss.

KW Human immunodeficiency virus type C strain AF110967.

OS Synthetic.

XX WO2000039304-A2.

XX 06-JUL-2000.

XX 30-DEC-1999; 94WO-US31273.

XX 31-DEC-1998; 94NS-0114495.

XX 01-SEP-1999; 94US-0152195.

XX (CHIR) CHIRON CORP.

XX Barnett S, Zur Megede J;

XX WPI: 2000-452401/39.

XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV

PT Env polypeptide and the polypeptide useful for immunizing a mammal

PT especially human against HIV

XX

PS Disclosure; Page 93; 113pp; English

XX Expression cassettes comprising a polynucleotide encoding antigenic

CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are

CC useful in DNA immunization, generation of packaging cell lines and

CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag

CC expression cassettes exhibit increased potency for induction of

CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1

CC self-assemble into non-infectious virus-like particles which are used as

CC a matrix for the proper presentation of an antigen entrapped or

CC associated to the immune system of the host.

XX Sequence 1509 BP; 320 A; 556 C; 472 G; 161 T; 0 other;

SQ

Query Match 100.0%; Score 60; DB 21; Length 1509;

Best Local Similarity 100.0%; Pred. No. 9e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatccgagggcccaaggagcccttcgcgactacgtgagccgtttcttcaagacc 60

DB 841 gacatccgagggcccaaggagcccttcgcgactacgtgagccgtttcttcaagacc 900

RESULT 3

AAA51626

ID AAA51626 standard; DNA; 1509 BP.

AC AAA51626;

XX 31-OCT-2000 (first entry)

XX HIV codon-optimized synthetic Gag polynucleotide.

DE

XX Gag; expression cassette; antigenic, type C; HIV; Env; synthetic;

KW DNA immunization; packaging cell line; antigen presentation; ss.

KW Human immunodeficiency virus type C strain AF110967.

OS Synthetic.

XX WO2000049404-A2.

XX 06-JUL-2000.

XX 30-DEC-1999; 94WO-US31273.

XX 31-DEC-1998; 94US-0114495.

XX 01-SEP-1999; 94US-0152195.

XX (CHIR) CHIRON CORP.

XX Barnett S, Zur Megede J;

XX WPI: 2000-452401/39.

XX Polynucleotide encoding antigenic type C HIV Gag polypeptide or a HIV

PT Env polypeptide and the polypeptide useful for immunizing a mammal

PT especially human against HIV

XX

PS Disclosure; Page 104; 113pp; English.

XX Expression cassettes comprising a polynucleotide encoding antigenic

CC type C human immunodeficiency virus (HIV) Gag or Env polypeptides are

CC useful in DNA immunization, generation of packaging cell lines and

CC production of Gag- and/or Env-containing proteins. Synthetic Env and Gag

CC expression cassettes exhibit increased potency for induction of

CC cytotoxic T-lymphocyte (CTL) responses by DNA immunization. Gag of HIV-1

CC self-assemble into non-infectious virus-like particles which are used as

CC a matrix for the proper presentation of an antigen entrapped or

CC associated to the immune system of the host.

XX Sequence 1509 BP; 321 A; 559 C; 471 G; 158 T; 0 other;

SQ

Query Match 100.0%; Score 60; DB 21; Length 1509;

Best Local Similarity 100.0%; Pred. No. 9e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatccgagggcccaaggagcccttcgcgactacgtgagccgtttcttcaagacc 60

DB 841 gacatccgagggcccaaggagcccttcgcgactacgtgagccgtttcttcaagacc 900

RESULT 4

AAA70426

ID AAA70426 standard; DNA; 60 BP.

XX


```

XX 31-DEC-1998; 98US-0114495.
XX 01-DEC-1999; 99US-0168471.
XX (CHIR ) CHIRON CORP.
XX Barnett S, Zur Meqede J, Srivastava I, Lian Y, Hartog K, Liu H,
XX Greer C, Selby M, Walker C;
XX WPI: 2000-452400/39.
XX
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
XX Gag-containing polypeptide useful for vaccinating against HIV
XX infections and acquired immunodeficiency syndrome (AIDS) -
XX
XX Claim 3; Fig 7; 391pp, English.
XX
XX The present sequence is the coding sequence of a HIV Gag expression
XX cassette, Gag.MuSR2. The Gag protein of HIV is needed for the assembly
XX of virus-like particles. In addition, the Gag protein is involved in
XX many stages of the HIV life cycle, including assembly, virion maturation
XX after particle release and early post-entry steps in viral replication.
XX The expression cassette may be used for the recombinant expression of
XX HIV Gag-polypeptides which may then be used to vaccinate against HIV
XX infection and acquired immunodeficiency syndrome (AIDS) .
XX
XX Sequence 1515 BP; 429 A; 547 C; 480 G; 159 T, 0 other;
XX
XX Query Match 97.3%; Score 58.4; DB 21; Length 1515;
XX Best local Similarity 68.3%; P1-J No. 2 7a-09;
XX Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps
XX
XX 1 gacatccgcaggcccccagaagagcccttcgcgcactactcgtggacccgtttctcaagacc 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 862 gacatccgcaggcccccagaagagcccttcgcgcactactcgtggacccgtttctcaagacc 92
XX
XX RESULT 8
XX AAA70413
XX ID AAA70413 standard; DNA; 1853 BP.
XX
XX AAC
XX AAA70413;
XX
XX 28-NOV-2000 (first entry)
XX
XX HIV Gag-protease expression cassette coding sequence GagProt ModS.
XX HIV-1; AIDS; Gag protease; vaccine; expression cassette; ss.
XX
XX Human immunodeficiency virus type 1.
XX Synthetic.
XX
XX WO200049302-A2.
XX
XX 06-JUL-2000.
XX
XX 30-DEC-1999; 99WO-US31245.
XX
XX 31-DEC-1998; 98US-0114495.
XX 01-DEC-1999; 99US-0168471.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Barnett S, Zur Meqede J, Srivastava I, Lian Y, Hartog K, Liu H,
XX Greer C, Selby M, Walker C;
XX WPI: 2000-452400/39.
XX
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
XX Gag-containing polypeptide useful for vaccinating against HIV
XX infections and acquired immunodeficiency syndrome (AIDS) -
XX
XX

```

PS Claim 5; Fig 7; 61pp; English.

XX The present sequence is the coding sequence of a HIV GAG protease
 CC expression cassette, GADPOT-ModS. The GAG protein of HIV is needed for
 CC the assembly of virus-like particles. In addition, the GAG protein is
 CC involved in many stages of the HIV life cycle, including assembly, virion
 CC maturation after particle release and early post entry steps in viral
 CC replication. The expression cassette may be used for the recombinant
 CC expression of HIV GAG polypeptides which may then be used to vaccinate
 CC against HIV infection and acquired immunodeficiency syndrome (AIDS).
 XX
 SQ Sequence 1865 BP; 421 A; 624 G; 589 C; 228 T; 0 other;

Query Match 97.38; Score 58.4; DB 21; Length 1865;
 Best Local Similarity 98.38; Pred. No. 2,86-09;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 gaatcccaacaccccaagagacccctccgaactactgaagacccctctcaaacac 60
 |||||||
 DB 862 gaatcccaacaccccaagagacccctccgaactactgaagacccctctcaaacac 921

RESULT 9
 AAA70475
 ID AAA70475 standard; DNA; 1865 BP.

XX AAA70475;

XX 28-NOV-2000 (first entry)

DE HIV GAG protease expression cassette coding sequence GADPOT-Mod.SF2(GP1).

KW HIV 1; AIDS; GAG protease; vaccine; expression cassette; ss.

XX Human immunodeficiency virus type 1.
 OS Synthetic.

XX W0200049492-A2.

XX 06-JUL-2000.

XX 30-DEC-1999; 99WO-US01245.

XX 31-DEC-1999; 98US-0114495.

XX 01-DEC-1999; 99US-0168471.

XX (CHIR) CHIRON CORP.

XX Barnett S., Zur Megede J., Srivastava L., Lian Y., Harlow K., Liu H;

XX Greer C., Selby M., Walker G;

XX W01; 2000 452409/39.

XX Expression cassettes encoding the human immunodeficiency virus (HIV)

XX GAG containing polypeptide useful for vaccinating against HIV

XX infections and acquired immunodeficiency syndrome (AIDS) -

XX Claim 5; Fig 69; 39pp; English

XX The present sequence is the coding sequence of a HIV GAG protease
 CC expression cassette, GADPOT-Mod.SF2 (GP1). The GAG protein of HIV is
 CC needed for the assembly of virus-like particles. In addition, the GAG
 CC protein is involved in many stages of the HIV life cycle, including
 CC assembly, virion maturation after particle release and early post-entry
 CC steps in viral replication. The expression cassette may be used for the
 CC recombinant expression of HIV GAG polypeptides which may then be used to
 CC vaccinate against HIV infection and acquired immunodeficiency syndrome
 CC (AIDS).

XX Sequence 1865 BP; 424 A; 627 G; 583 C; 231 T; 0 other;

Query Match 97.38; Score 58.4; DB 21; Length 1865;
 Best Local Similarity 98.38; Pred. No. 2,86-09;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 gaatcccaacaccccaagagacccctccgaactactgaagacccctctcaaacac 60
 |||||||
 DB 868 gaatcccaacaccccaagagacccctccgaactactgaagacccctctcaaacac 927

RESULT 10
 AAA70476
 ID AAA70476 standard; DNA; 1865 BP.

XX AAA70476;

XX 28-NOV-2000 (first entry)

DE HIV GAG protease expression cassette coding sequence GADPOT-Mod.SF2(GP2).

KW HIV 1; AIDS; GAG protease; vaccine; expression cassette; ss.

XX Human immunodeficiency virus type 1.

OS Synthetic.

XX W0200039402-A2.

XX 06-JUL-2000.

XX 30-DEC-1999; 99WO-US01245.

XX 31-DEC-1999; 98US-0114495.

XX 01-DEC-1999; 99US-0168471.

XX (CHIR) CHIRON CORP.

XX Barnett S., Zur Megede J., Srivastava L., Lian Y., Harlow K., Liu H;

XX Greer C., Selby M., Walker G;

XX W01; 2000 452409/39.

XX Expression cassettes encoding the human immunodeficiency virus (HIV)

XX GAG-containing polypeptide useful for vaccinating against HIV

XX infections and acquired immunodeficiency syndrome (AIDS) -

XX Claim 5; Fig 70; 39pp; English.

XX The present sequence is the coding sequence of a HIV GAG protease
 CC expression cassette, GADPOT-Mod.SF2(GP2). The GAG protein of HIV is
 CC needed for the assembly of virus-like particles. In addition, the GAG
 CC protein is involved in many stages of the HIV life cycle, including
 CC assembly, virion maturation after particle release and early post-entry
 CC steps in viral replication. The expression cassette may be used for the
 CC recombinant expression of HIV GAG polypeptides which may then be used to
 CC vaccinate against HIV infection and acquired immunodeficiency syndrome
 CC (AIDS).

XX Sequence 1865 BP; 460 A; 583 C; 569 G; 253 T; 0 other;

Query Match 97.38; Score 58.4; DB 21; Length 1865;
 Best Local Similarity 98.38; Pred. No. 2,86-09;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 gaatcccaacaccccaagagacccctccgaactactgaagacccctctcaaacac 60
 |||||||
 DB 862 gaatcccaacaccccaagagacccctccgaactactgaagacccctctcaaacac 927

RESULT 11
 AAA70415
 ID AAA70415 standard; DNA; 2031 BP.

XX AAA70415;

```

PD   06-JUL-2000..
XX
XX   30-DEC-1999;    99WD-US31245.
XX PF
XX   31-DEC-1998;    98US-0114495.
XX PR
XX   01-DEC-1999;    99US-0168471.
XX PP
XX (CHIR ) CHIRON CORP.
XX PA
XX Barnett S., Zur Megede J., Srivastava I., Lian Y., Hartog K., Liu H.;
XX PI Greer C., Selby M., Walker C.;
XX XX
XX WPI: 2000-452400/39.
XX DR
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
XX PT Gag-containing polypeptide useful for vaccinating against HIV
XX PT infections and acquired immunodeficiency syndrome (AIDS) -
XX PS
XX Claim 10; Fig 7; 39lpp; English.
XX XX
XX The present sequence is the coding sequence of a HIV Gag-polymerase
XX CC expression cassette, Gcspol ModSF. The Gag protein of HIV is needed for
XX CC the assembly of virus-like particles. In addition, the Gag protein is
XX CC involved in many stages of the HIV life cycle, including assembly, virion
XX CC maturation after particle release and early post entry steps in viral
XX CC replication. The expression cassette may be used for the recombinant
XX CC expression of HIV Gag polypeptides which may then be used to vaccinate
XX CC against HIV infection and acquired immunodeficiency syndrome (AIDS).
XX XX
XX Sequence 4319 BP; 1003 A; 1426 C; 1361 G; 529 T; 0 other;
XX SQ

Query Match          97.3%; Score 58.4; DB 21; Length 4319;
Best Local Similarity 98.3%; Pred. No. 3e-09;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q#      1 gacatccgcagcgggcccccaaggagcccttcgcgactcagtgaccgcccttacaagacc 60
Db       |||||||
862 gacatccgcagcgcccgaaggagcccttcgcgactcagtgaccgcccttacaagacc 921

RESULT 13
AAAT0472
ID     AAAT0472 standard; DNA; 4472 BP.
XX
XX     AAA'0472;
XX AC
XX XX
XX NT 28-Nov-2000 (first entry)
XX DE HIV bicistronic construct g1603 modUS4 delV1/V2 Gag.modSF2.
XX KW
XX HIV-1; AIDS; Gag; vaccine; expression cassette; Env; SS.
XX XX
XX Chimeric - Cytomegalovirus
XX QS Chimeric - Human immunodeficiency virus type 1.
XX XX
XX PN WO200039302-A2.
XX XX
XX 06-JUL-2000..
XX XX
XX 30-DEC-1999;    99WD-US31245.
XX PF
XX 31-DEC-1998;    98US-0114495.
XX PP 01-DEC-1999;    99US-0168471.
XX XX
XX (CHIR ) CHIRON CORP.
XX PA
XX Barnett S., Zur Megede J., Srivastava I., Lian Y., Hartog K., Liu H.;
XX PI Greer C., Selby M., Walker C.;
XX XX
XX WPI: 2000-452400/39.
XX DR
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
XX PT

```


Query Match 97.3%; Score 58.4; DB 21; length 4689;
 Best Local Similarity 98.3%; Pred. No. 3e-09;
 Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 gacatccgcccaaggccccaaggagagcccttcgcggaactacgtagacgcgtttctcaagacc 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 4011 gacatccgcccaaggccccaaggagagcccttcgcggaactacgtagacgcgtttctcaagacc 4070

Search completed: August 8, 2001, 18:25:01
 Job time: 12605 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

em nucleic - nucleic search, using sw model

Run on: August 8, 2001, 20:00:52 ; Search time 5719.75 Seconds
(without alignments)
162.256 Million cell updates/sec

Title: US-09-475-704-2
Percent score: 60
Sequence: 1 qacatccqacqacccccc.....tgaacgcttcttcaaac 60

Scoring table: IDENTITY_NBC
Gapop 10.0 , Gapext 1.0

Searched: 144157 seqs, 773874588 residues 2688314
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenBank : *
- 1: qb_ba1:*
 - 2: qb_ba2:*
 - 3: qb_ba3:*
 - 4: qb_in1:*
 - 5: qb_in2:*
 - 6: qb_in3:*
 - 7: qb_em:*
 - 8: qb_ov:*
 - 9: qb_pat1:*
 - 10: qb_pat2:*
 - 11: qb_ph:*
 - 12: qb_pl1:*
 - 13: qb_pl2:*
 - 14: qb_pl3:*
 - 15: qb_pl4:*
 - 16: em_ba1:*
 - 17: em_ba2:*
 - 18: em_fun:*
 - 19: em_hqo_hum:*
 - 20: em_hqo_inv:*
 - 21: em_hqo_rnd:*
 - 22: em_hq_hum1:*
 - 23: em_hq_hum2:*
 - 24: em_hq_hum3:*
 - 25: em_hq_hum4:*
 - 26: em_hq_hum5:*
 - 27: em_hq_hum6:*
 - 28: em_hq_hum7:*
 - 29: em_hq_hum8:*
 - 30: em_hq_inv1:*
 - 31: em_hq_inv2:*
 - 32: em_hq_other:*
 - 33: em_hq_rnd:*
 - 34: em_hum1:*
 - 35: em_hum2:*
 - 36: em_hum3:*
 - 37: em_hum4:*
 - 38: em_hum5:*
 - 39: em_hum6:*
 - 40: em_hum7:*
 - 41: em_in:*
 - 42: em_em:*
 - 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_v1:*
- 53: qb_sts1:*
- 54: qb_sts2:*
- 55: qb_sts3:*
- 56: qb_sy:*
- 57: qb_un:*
- 58: qb_v11:*
- 59: qb_v12:*
- 60: qb_hq1:*
- 61: qb_hq2:*
- 62: qb_hq3:*
- 63: qb_hq4:*
- 64: qb_hq5:*
- 65: qb_hq6:*
- 66: qb_hq7:*
- 67: qb_hq8:*
- 68: qb_hq9:*
- 69: qb_hq10:*
- 70: qb_hq11:*
- 71: qb_hq12:*
- 72: qb_hq13:*
- 73: qb_hq14:*
- 74: qb_hq15:*
- 75: qb_hq16:*
- 76: qb_hq17:*
- 77: qb_hq18:*
- 78: qb_hq19:*
- 79: qb_hq20:*
- 80: qb_hq21:*
- 81: qb_hq22:*
- 82: qb_hq23:*
- 83: qb_hq24:*
- 84: qb_hq25:*
- 85: qb_pr1:*
- 86: qb_pr2:*
- 87: qb_pr3:*
- 88: qb_pr4:*
- 89: qb_pr5:*
- 90: qb_pr6:*
- 91: qb_pr7:*
- 92: qb_pr8:*
- 93: qb_pr9:*
- 94: qb_rnd:*
- 95: qb_rnd2:*
- 96: qb_in4:*
- 97: qb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB ID	Description
	Score	Match	Length		
1	58.4	97.3	1509	56	AF201927 Synthetic
2	58.4	97.3	1917	56	AF202464 Synthetic
3	58.4	97.3	1847	56	AF202465 Synthetic
4	48.8	81.3	1548	56	AF287354 Synthetic
5	48.8	81.3	4352	56	AF287352 Synthetic
6	48.8	81.3	4353	56	AF287353 Synthetic
7	47.8	79.7	4307	9	AX019132 Sequence
8	47.8	79.7	4307	9	AX035453 Sequence

```

9 47.8 79.7 4307 9 AX050836 Sequence
10 47.8 79.7 4327 9 AX035465 Sequence
11 47.8 79.7 4353 9 AX035464 Sequence
12 47.8 79.7 4353 9 AX035464 Sequence
13 38.2 63.7 693 58 AB034323 Human imm
14 38.2 63.7 9010 58 AF110974 HIV-1 iso
15 38.2 63.7 9010 58 AF110975 HIV-1 iso
16 38.2 63.7 9012 58 AF110973 HIV-1 iso
17 36.6 61.0 570 59 HIV-2-4196 Human imm
18 36.6 61.0 646 59 HIV-2-4196 Human imm
19 36.6 61.0 694 58 AB044421 Human imm
20 36.6 61.0 693 58 AB044422 Human imm
21 36.6 61.0 693 58 AB034324 Human imm
22 36.6 61.0 693 58 AB034325 Human imm
23 36.6 61.0 693 58 AB034326 Human imm
24 36.6 61.0 693 58 AB034327 Human imm
25 36.6 61.0 8975 59 U88822 HIV-1 iso
26 36.6 61.0 9034 58 AF286234 HIV-1 str
27 36.6 61.0 9811 58 CIVCG X52154 Chimpanzee
28 36.6 61.0 663 59 HIV274572 Human imm
29 36.6 61.0 1015 59 HIV1B2163R Human imm
30 36.6 61.0 9012 58 AF286235 HIV-1 str
31 35 58.3 313 58 AF319334 HIV-1 iso
32 35 58.3 313 58 AF319347 HIV-1 iso
33 35 58.3 313 58 AF319356 HIV-1 iso
34 35 58.3 313 58 AF319359 HIV-1 iso
35 35 58.3 313 58 AF319378 HIV-1 iso
36 35 58.3 313 58 AF319379 HIV-1 iso
37 35 58.3 389 59 HIV073073 HIV-1 iso
38 35 58.3 393 58 AF106450 HIV-1 iso
39 35 58.3 393 58 AF106451 HIV-1 iso
40 35 58.3 393 58 AF106453 HIV-1 iso
41 35 58.3 424 58 AF184568 HIV-1 iso
42 35 58.3 468 58 AF260499 HIV-1 iso
43 35 58.3 489 58 AF184521 HIV-1 iso
44 35 58.3 587 58 AF184520 HIV-1 iso
45 35 58.3 587 59 HIV274541 Human imm

```

ALIGNMENTS

```

RESULT 1
AF201927 1509 bp DNA 15-MAP-2000
LOCUS Synthetic construct gag protein gene, complete cds.
ACCESSION AF201927
VERSION AF201927.1 GI:7248702
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
ARTIFICIAL SEQUENCE.
REFERENCE 1 (bases 1 to 1509)
AUTHORS zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
TITLE Increased expression and immunogenicity of sequence-modified human
immunodeficiency virus type 1 gag gene
JOURNAL J. Virol. 74 (6), 2628-2635 (2000)
MEDLINE 20148954
PUBMED 10684277
REFERENCE 2 (bases 1 to 1509)
AUTHORS zur Megede,J. and Barnett,S.W.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton,
Emeryville, CA 94608, USA
FEATURES
source 1. 1509
/db_xref="taxon:32630"
/organism="synthetic construct"
CDS
1..1509
/translation="MGARASVLSGGELDKWKIRLPGGKKYKIKLHIVASRELRF
AVNPGLETSEGGKQILGGVLSQGTSEELSLYNTVATLYCVHORIDVKDTEALE
KIEEONKSKRAQAIAAAAGTGNSSQVNPPIVONLOGVMHQAISPTLNAMKVY
VEEKAFSPVIMFSALESAGTQDLNTMLNTVGHGAAMOMLKETINEEAEDRVH
IYVAGTATTCQMPREPCSDIAGTSTLQEQICWMNNPPIPVGEIYKWIILGLNKIV
PMYSGTILDPQCFKPFPPFVDFPYKTLPAFAQSQVKNMWTETILVGNANDCKT
ILKALGPAATLEEMTACQGVGCHIKARVLAZAMSQVTFATIMMGKGFNPKQKTV
KCFNGKEGHTARNCPAPRKKCPWPGPCHMKPTEPCANFLGTPWSYKPPGPF
LQSRPEPTAPPESPPFPEKTPSQKQRPIDKILYPLTSLKSLGNDPSSQ"
BASE COUNT 420 a 620 c 579 g 228 t
ORIGIN
Query Match 97.3% Score 58.4; DB 56; Length 1847;

```

```

/product="gag protein"
/protein_id="AAF43628.1"
/db_xref="GI:7248703"
/translation="MGAPASVLSGCELKWKIKPLPPGKKYKIKHIVASRELFFF
AVNGLLETSEGGKQILGQLFSLQGTSEELSLYNTVATLYCVHORIDVKDTEALE
KIEEONKSKRAQAIAAAAGTGNSSQVNPPIVONLOGVMHQAISPTLNAMKVY
VEEKAFSPVIMFSALESAGTQDLNTMLNTVGHGAAMOMLKETINEEAEDRVH
IYVAGTATTCQMPREPCSDIAGTSTLQEQICWMNNPPIPVGEIYKWIILGLNKIV
PMYSGTILDPQCFKPFPPFVDFPYKTLPAFAQSQVKNMWTETILVGNANDCKT
ILKALGPAATLEEMTACQGVGCHIKARVLAZAMSQVTFATIMMGKGFNPKQKTV
KCFNGKEGHTARNCPAPRKKCPWPGPCHMKPTEPCANFLGTPWSYKPPGPF
LQSRPEPTAPPESPPFPEKTPSQKQRPIDKILYPLTSLKSLGNDPSSQ"
BASE COUNT 328 a 543 c 479 g 159 t
ORIGIN
Query Match 97.3% Score 58.4; DB 56; Length 1509;
Rest local Similarity 98.3% Pred No. 6.2e-07;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 qacatccacccagggccccagagaccccttcacgacactacgtgagccqctttcgaagacc 60
|||||
Db 856 GACATCGCGCAGGAGCCCAAGGAGCCCTTCGCGGACTAGTGAGCGCTTTTACAGAGCC 915
|||||
RESULT 2
AF202464 1847 bp DNA SYN 12-MAP-2000
LOCUS Synthetic construct gag-protease fusion protein gpi gene, complete
DEFINITION cds.
ACCESSION AF202464
VERSION AF202464.1 GI:7229424
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
ARTIFICIAL SEQUENCE.
REFERENCE 1 (bases 1 to 1847)
AUTHORS zur Megede,J., Chen,M.C., Doe,B., Schaefer,M., Greer,C.E.,
Selby,M., Otten,G.R. and Barnett,S.W.
TITLE Increased expression and immunogenicity of sequence-modified human
immunodeficiency virus type 1 gag gene
JOURNAL J. Virol. 74 (6), 2628-2635 (2000)
MEDLINE 20148954
PUBMED 10684277
REFERENCE 2 (bases 1 to 1847)
AUTHORS zur Megede,J. and Barnett,S.W.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1999) Vaccines, Chiron Corporation, 4560 Horton
Street, Emeryville, CA 94608, USA
FEATURES
source 1..1847
/organism="synthetic construct"
/db_xref="taxon:32630"
1..1509
/transl_table=11
/product="gag-protease fusion protein gpi"
/protein_id="AAF42819.1"
/db_xref="GI:7229425"
/translation="MGARASVLSGGELDKWKIRLPGGKKYKIKLHIVASRELRF
AVNPGLETSEGGKQILGGVLSQGTSEELSLYNTVATLYCVHORIDVKDTEALE
KIEEONKSKRAQAIAAAAGTGNSSQVNPPIVONLOGVMHQAISPTLNAMKVY
VEEKAFSPVIMFSALESAGTQDLNTMLNTVGHGAAMOMLKETINEEAEDRVH
IYVAGTATTCQMPREPCSDIAGTSTLQEQICWMNNPPIPVGEIYKWIILGLNKIV
PMYSGTILDPQCFKPFPPFVDFPYKTLPAFAQSQVKNMWTETILVGNANDCKT
ILKALGPAATLEEMTACQGVGCHIKARVLAZAMSQVTFATIMMGKGFNPKQKTV
KCFNGKEGHTARNCPAPRKKCPWPGPCHMKPTEPCANFLGTPWSYKPPGPF
LQSRPEPTAPPESPPFPEKTPSQKQRPIDKILYPLTSLKSLGNDPSSQ"
BASE COUNT 420 a 620 c 579 g 228 t
ORIGIN
Query Match 97.3% Score 58.4; DB 56; Length 1847;

```



```

source
1. 4353
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="derived from Human Immunodeficiency virus type 1 isolate YU-2"
37. 4344
/note="codon optimized reading frame"
/codon_start=1
/transl_table=11
/product="gay-poi fusion protein"
/protein_id="AAG28733.1"
/db_xref="GI:11066865"
<338..4343
/note="codon optimized reading frame"
/codon_start=1
/transl_table=11
/product="pol protein"
/protein_id="pol protein"
/db_xref="GI:11066863"
/transl_table=11
/note="BED:APPOCKARKTSSEGTIPANSPIPEPQWVHNNNSLSEAGA
DROGTVSFPOITIMORPLVTIKGQIKKALDITGADDTVLEEMNLPGRWPKMIG
GIGGFVKVQYDQIPTEIGCHRAIKTALVIGTPVNIIGKLLTQIGCTLNPSPLET
VPKLPKMGDPKVKOMPLTEKIKALVEICTEMEKEGKISKIGPFNPYPVFAIKK
KSTWKPTLVPFPLNKRTPQWVQVQIGIPHAGIKKKSVTVLVDGDAYEVSPLHED
FRAYTPTLPSINNETPGTYOYNYVPOGKGPALFOSSMTTIIPEPKONPLDITY
QYMDIIVSDEIGQHRKIEBLRHLLKWFUTPKKHOKPEPLMKGVLHPDKW
TQVPIVPERKSDVNDIQKLVNWSQILAEIKQKQGVQYQYQYQEPFNKTKGKA
EAELEARENILKEPVHYYDPKSDLIAEIKQKQGVQYQYQYQEPFNKTKGKA
RTGANTNDVKLTQAVKIAIESTVIMSKTPFKTOKETWEITWETWQATWIE
WEVNTPLVKKWLTQVKEPIIGAEYFYDGAANRETKLGRAGYVTKGKQVSLTD
TTNKPLQAGIYALQDSLEVINIVLSVAGILQAGURSESELSVQLLEQLIKKE
KYLAVAPKAILGGNPNQVNVLSACIPKVLFDGDKAOPHEKYNHNSPAMASDN
LPPVAKELVASCDKQLKGAMHGVNCSHILWQDCHLGGKVLIVAVHVASGYIE
AEVIAETGQETAYFLKLAGRPVTTIHTDGNSTPATSATVKAACWAGIKQFGPIY
NPQSGVSEMNKELKIIIGVQVQAEHLKTAQWAVFTHNFKRKGIGRYSAGERIV
DIATDITQIKELQKQIKQNFVYVYRISHIPKTPARKLLNWSGEAYVIQDSIKV
VPRKAKIIRDYGKUMAGDUCVAGRODED"
BASE COUNT 1066 a 1344 c 1410 g 532 t
ORIGIN

Query Match 81.3% Score 48.8; DB 56; Length 4352;
Best Local Similarity 88.3%; Pred No 0.00023;
Matches 53; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gacatccgcagggcccaagagagcccttcgcgcactacatgagagcgtctctcagagac 60
||||| 1 ||||||||||||||||||| 1 ||||||||||||||| 1 ||||| 1 |||||
DB 886 GACATCAGGCAAGAGGACCAAGAGAGGACGACGACGACGACGACGACGACGAC 945

RESULT 7
LOCUS AX019132 4307 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 2 from Patent W09941397.
ACCESSION AX019132
VERSION AX019132.1 GI:10043165
KEYWORDS Human immunodeficiency virus.
SOURCE Human immunodeficiency virus.
ORGANISM Human immunodeficiency virus.
VIRUSES: Retroviridae; Retroviridae; Lentivirus; Primate
lentivirus group.
1 (bases 1 to 4307)
Mitrepharous, K., Kingsman, A.J. and Kim, N.
Anti-viral vectors
TITLE Anti-viral vectors
JOURNAL Patent: WO 9941397-A 2 19-AUG-1999;
BIOMEDICA LTD (GB); KIM NARRY (KR)
LOCATION/Qualifiers
1. 4307
/organism="Human immunodeficiency virus"
/db_xref="taxon:12721"
BASE COUNT 1114 a 1214 c 1344 g 685 t
ORIGIN

Query Match 81.3% Score 48.8; DB 56; Length 4352;
Best Local Similarity 88.3%; Pred No 0.00023;
Matches 53; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gacatccgcagggcccaagagagcccttcgcgcactacatgagagcgtctctcagagac 60
||||| 1 ||||||||||||||||||| 1 ||||||||||||||| 1 ||||| 1 |||||
DB 886 GACATCAGGCAAGAGGACCAAGAGAGGACGACGACGACGACGACGACGACGAC 945

RESULT 6
LOCUS AF287353 4353 bp DNA SYN 01-NOV-2000
DEFINITION Synthetic construct HIV-1-derived gay-poi fusion protein gene,
complete cds.
ACCESSION AF287353
VERSION AF287353.1 GI:11066864
KEYWORDS gay-poi fusion protein.
SOURCE gay-poi fusion protein.
ORGANISM gay-poi fusion protein.
ARTIFICIAL SEQUENCE.
1 (bases 1 to 4353)
Fuller, M. and Anson, D.S.
TITLE Helper plasmids for production of HIV derived vectors
JOURNAL Unpublished
2 (bases 1 to 4353)
Fuller, M. and Anson, D.S.
TITLE Direct Submission
JOURNAL Submitted (12-Jul-2000) National Pathology, Women's and Children's
Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
LOCATION/Qualifiers

```

```

Query Match          79.7%; Score 47.8; DB 9; Length 4307;
Best local Similarity 88.1%; Pred. No. 0.00045;
Matches 52; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 qacatccgcacagggccggaaggaacgcttttcgggactagctagcagcgtttctcaaac 59
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 850 GACATCGGCGAAGGCGGGAAGGAACGCTTTTCGGGACTAGCTAGCGGCTTCTACAAAAC 908

RESULT 10
AX035465
LOCUS AX035465 4307 bp DNA PAT 15-NOV-2000
DEFINITION Sequence 2 from Patent W00055341.
ACCESSION AX035465
VERSION AX035465.1 GI:11191095
KEYWORDS Synthetic construct.
SOURCE Synthetic construct.
ORGANISM Artificial sequence.
REFERENCE 1 (bases 1 to 4307)
AUTHORS Uden,M. and Mitrophanous,K.
TITLE Anti-viral vectors
JOURNAL Patent: WO 0055341-A 2 21-SEP-2000.
UDEN MARK (GB) : OXFORD BIOMEDICA LTD (GB) ; MITROPHANOUS KYRIACOS
(US)

FEATURES
    source
        Location/Qualifiers
            1..4307
                /organism="Synthetic construct"
                /db_xref="taxon:32630"
                /note="haplo-SYNGM - codon optimised gagpol sequence"

BASE COUNT 1135 a 1198 c 1285 g 689 t

ORIGIN
Query Match          79.7%; Score 47.8; DB 9; Length 4307;
Best local Similarity 88.1%; Pred. No. 0.00045;
Matches 52; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 qacatccgcacagggccggaaggaacgcttttcgggactagctagcagcgtttctcaaac 59
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 850 GACATCGGCGAAGGCGGGAAGGAACGCTTTTCGGGACTAGCTAGCGGCTTCTACAAAAC 908

RESULT 9
AX035466
LOCUS AX035466 4307 bp DNA PAT 17-JAN-2001
DEFINITION Sequence 2 from Patent W00075370.
ACCESSION AX035466
VERSION AX035466.1 GI:12409774
KEYWORDS Synthetic construct.
SOURCE Synthetic construct.
ORGANISM Artificial sequence.
REFERENCE 1 (bases 1 to 4307)
AUTHORS Mitrophanous,K., Kim,N.H. and Kotsopoulos,E.
TITLE In vivo selection method for determining inhibitory rna molecules
JOURNAL Patent: WO 0075370-A 2 14-DEC-2000;
OXFORD BIOMEDICA (UK) LIMITED (GB)

FEATURES
    source
        Location/Qualifiers
            1..4307
                /organism="Synthetic construct"
                /db_xref="taxon:32630"
                /note="Codon optimised gagpol sequence"

BASE COUNT 1135 a 1198 c 1285 g 689 t

ORIGIN
Query Match          79.7%; Score 47.8; DB 9; Length 4307;
Best local Similarity 88.1%; Pred. No. 0.00045;
Matches 52; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 qacatccgcacagggccggaaggaacgcttttcgggactagctagcagcgtttctcaaac 59
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 850 GACATCGGCGAAGGCGGGAAGGAACGCTTTTCGGGACTAGCTAGCGGCTTCTACAAAAC 908

RESULT 10
AX035465
LOCUS AX035465 4327 bp DNA PAT 15-NOV-2000
DEFINITION Sequence 14 from Patent W00055341.
ACCESSION AX035465
VERSION AX035465.1 GI:11191107
KEYWORDS Synthetic construct.
SOURCE Synthetic construct.
ORGANISM Artificial sequence.
REFERENCE 1 (bases 1 to 4327)
AUTHORS Uden,M. and Mitrophanous,K.
TITLE Anti-viral vectors
JOURNAL Patent: WO 0055341-A 14 21-SEP-2000;
UDEN MARK (GB) : OXFORD BIOMEDICA LTD (GB) ; MITROPHANOUS KYRIACOS
(US)

FEATURES
    source
        Location/Qualifiers
            1..4327
                /organism="Synthetic construct"
                /db_xref="taxon:32630"
                /note="SYNGM - codon optimised HIV-1 gagpol with 20 bp
of the leader sequence of HIV-1, etc"

BASE COUNT 1144 a 1194 c 1296 g 693 t

ORIGIN
Query Match          79.7%; Score 47.8; DB 9; Length 4327;
Best local Similarity 88.1%; Pred. No. 0.00045;
Matches 52; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 qacatccgcacagggccggaaggaacgcttttcgggactagctagcagcgtttctcaaac 59
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 870 GACATCGGCGAAGGCGGGAAGGAACGCTTTTCGGGACTAGCTAGCGGCTTCTACAAAAC 928

RESULT 11
AX035464
LOCUS AX035464 4353 bp DNA PAT 15-NOV-2000
DEFINITION Sequence 13 from Patent W00055341.
ACCESSION AX035464
VERSION AX035464.1 GI:11191106
KEYWORDS Synthetic construct.
SOURCE Synthetic construct.
ORGANISM Artificial sequence.
REFERENCE 1 (bases 1 to 4353)
AUTHORS Uden,M. and Mitrophanous,K.
TITLE Anti-viral vectors
JOURNAL Patent: WO 0055341-A 13 21-SEP-2000;
UDEN MARK (GB) : OXFORD BIOMEDICA LTD (GB) ; MITROPHANOUS KYRIACOS
(US)

FEATURES
    source
        Location/Qualifiers
            1..4353
                /organism="Synthetic construct"
                /db_xref="taxon:32630"
                /note="SYNGM - codon optimised HIV-1 gagpol with leader
sequence from the major splice donor"

BASE COUNT 1153 a 1198 c 1302 g 700 t

ORIGIN
Query Match          79.7%; Score 47.8; DB 9; Length 4353;
Best local Similarity 88.1%; Pred. No. 0.00045;
Matches 52; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 qacatccgcacagggccggaaggaacgcttttcgggactagctagcagcgtttctcaaac 59
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 896 GACATCGGCGAAGGCGGGAAGGAACGCTTTTCGGGACTAGCTAGCGGCTTCTACAAAAC 954

```

```

RESULT 12
AX035463 LOCUS AX035463 4642 bp DNA PAT 15-NOV-2000
DEFINITION Sequence 12 from Patent WO0055341.
ACCESSION AX035463
VERSION AX035463.1 GI:11191105
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 4642)
AUTHORS Uden,M. and Mitrophanous,K.
TITLE Anti-viral vectors
JOURNAL Patent. WO 6055341-A 12 21-SEP-2000;
UDEN MARK (GR) : OXFORD BIOMEDICA LTD (GB) : MITROPHANOUS KYRIACOS
(US)
FEATURES
source Location/Qualifiers
1..4642
/oranism="synthetic construct"
/db_xref="taxon:32630"
/notes="pcymer2 - codon optimised HIV-1 gagpol with leader
sequence"
BASE COUNT 1218 a 1273 c 1389 g 762 t
ORIGIN

Query Match 79.7%; Score 47.8; DB 9; Length 4642;
Best Local Similarity 98.1%; Pred. No. 0.00044;
Matches 52; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 qacatccgcagcccaagcagccctccgcagcactacatgacccctcttcacagac 59
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1185 GACATCCGCCCAAGGCGCTTATGAGATTTTGGAGATATGAGACGGTTCACAAAC 1243

RESULT 13
AB034323 LOCUS AB034323 693 bp DNA VRL 13-JUN 2000
DEFINITION Human immunodeficiency virus type 1 proviral gene for gag protein
p24 region, partial cDS, isolate-patient 5, clone:p24(5)-d.
ACCESSION AB034323
VERSION AB034323.1 GI:7416276
KEYWORDS Human immunodeficiency virus type 1 (isolate-patient 5) proviral
SOURCE DNA, clone:p24(5)-d.p24(5)-e.p24(5)-i.
ORGANISM Human immunodeficiency virus type 1
REFERENCE 1 (sites)
AUTHORS Yamada,T. and Iwamoto,A.
TITLE Comparison of proviral accessory genes between long term
nonprogressors and progressors of human immunodeficiency virus type
1 infection
JOURNAL Arch. Virol. 145 (5), 1021-1027 (2000)
MEDLINE 20338596
REFERENCE 2 (bases 1 to 693)
AUTHORS Yamada,T.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-1999) to the DDBJ/EMBL/GenBank databases. Takeshi
Yamada, Institute of Medical Science, University of Tokyo,
Department of Infectious Diseases; 4-6-1 Shirokanedai, Minato-ku,
Tokyo 108-8639, Japan (E-mail:takeshi.yamada@y.ims.u-tokyo.ac.jp,
Tel:81-3-5449-5336, Fax:81-3-5449-5427)
FEATURES
source Location/Qualifiers
1..693
/oranism="Human immunodeficiency virus type 1"
/proviral
/isolate="patient 5"
/db_xref="taxon:11676"
/clone="p24(5)-d.p24(5)-e.p24(5)-i"
1..693
gene

```

```

CDS
/gene="gag"
<1..>693
/gene="gag"
/note="p24 region"
/codon_start=1
/product="gag protein"
/protein_id="BAA93787.1"
/db_xref="GI:7416277"
/translation="PIVONLOGQWVHQAIQSPRLNANWKKVVEKAFSPVIMPFSALS
EGATPDQNLNTLVGGHQAAOMLKEITPEAAEDRLHIVHAGPIAGOMREPRGS
DIAGTITFLQIGIMTNINPPIVCEIYKRWIIGLNFIVMYSPITSIIIDPGPKP
PRVDVREYKTLRAFOASQEVKNMIEILLVQANPDKTILKALGPAATLEEMTAC
QGVGGPHKARVL"
BASE COUNT 249 a 140 c 169 g 135 t
ORIGIN

Query Match 63.7%; Score 38.2; DB 58; Length 693;
Best Local Similarity 78.0%; Pred. No. 0.37;
Matches 46; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 qacatccgcagcccaagcagccctccgcagcactacatgacccctcttcacagac 59
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 454 GACATAAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 512

RESULT 14
AF110974 LOCUS AF110974 9010 bp DNA VRL 02-MAR-2001
DEFINITION HIV-1 isolate C-96BWI5C02 country Botswana, complete genome.
ACCESSION AF110974
VERSION AF110974.1 GI:4324730
KEYWORDS Human immunodeficiency virus type 1.
SOURCE Human immunodeficiency virus type 1
ORGANISM Viruses; Retroviridae; Lentivirus; Primate
lentivirus group.
REFERENCE 1 (bases 1 to 9010)
AUTHORS Novitsky,V.A., Montano,M.A., McLane,M.F., Kenjifo,B., Vannberg,F.,
Foley,B.T., Ndung'u,T.P., Rahman,M., Makhema,M.J., Marlink,R. and
Essex,M.
TITLE Molecular cloning and phylogenetic analysis of human
immunodeficiency virus type 1 subtype C: a set of 23 full-length
clones from Botswana
JOURNAL J. Virol. 73 (5), 4427-4432 (1999)
MEDLINE 94214383
REFERENCE 10196340
AUTHORS Novitsky,V.A., Montano,M.A., McLane,M.F., Kenjifo,B., Vannberg,F.,
Foley,B.T., Ndung'u,T.P., Marlink,R. and Essex,M.
TITLE Direct Submission
JOURNAL Submitted (03-DEC 1998) Immunology and Infectious Diseases, Harvard
School of Public Health, FXB-310, 651 Huntington Avenue, Boston, MA
02115, USA
FEATURES
source Location/Qualifiers
1..9010
/oranism="Human immunodeficiency virus type 1"
/proviral
/isolate="C-96BWI5C02"
/db_xref="taxon:11676"
/country="Botswana"
/notes="Subtype: C"
1..124
281..4563
/gene="gag"
281..4563
/gene="pol"
join(281..1564,1564..4563)
/codon_start=1
/product="gag-pol polyprotein"
/protein_id="AA017126.1"
/db_xref="GI:4324862"
/translation "MCARASRLKJRKLD:WUKLRLFTGKKCYM:PHIVASPEIPPE

```



```

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
PFFPFFNF
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .9010
/organism="Human immunodeficiency virus type 1"
/proviral
/isolate="C-96RW15C05"
/db_xref="taxon:11676"
/country="Botswana"
/notes="subtype: C"
1. .124
281. .1759
/gene="gag"
281. .1759
/gene="pol"
join(281. .1564,1564 4563)
/codon_start=1
/product="gag-pol polyprotein"
/protein_id="AAD17135.1"
/db_xref="GI:4324871"
/translation="MGAKASILPGEKLDANERIPPPGKCYMYLKHIVWASRELERF
ALNFKLSEGGKQIVKQALQALQITGIELKSLNVAIYCVHEKIKVKDKKALD
KIEEONKSEKTKQAAAGAKTSQNPVQVQGMVHQAIISPTLNAAVVKVIEKA
FSPEVIMFMTALSEGATPDQNTMTNTVGGHQAAMKLDIINEEAAEWRLHPVHAG
PIACQKREPGSDIAGTSTLOEITMTSNPIIPVGIYKKWIILGLINKIVRMYSF
VSLIDIRQKREPRDYVDFKTLAEQATQDVKNMTDTLLVONANPCKTILRAL
GPATLEEMTACQVGGPGHKARVIAEAMSOATSNILMORSNFKGPKRIKFCNCG
KEGHIARNCRAPKKGCKGEGHQMCKTERQANFFREDLAFPOGKARFEPSEQR
ANSPTRELOVRGDNPRSEAGERTLNFQITLWOPLVSTIKVGGQIKKALDITGA
DDVLEMSLPGKWKPMIGGIGGFTKVRQYDQIIEJCGKKAIGTVIIGPTPPVNIIG
RNMLTQGLTNFPIPIETVPVKLPKMGPKVKOWPITEEIKATITACIEMKEG
KATLGPENYNTVPFAIKKDKKPKLVDEPFIKPTQDFWFOIGIPIHPAGLKKK
KSVTLVDGATFSVLDEDFRKYTAFTIPSINNETPGIRYOYNVLPOCWKSPSIEQ
SSMTKILPEFRANPELVYQMDLYVSDLEIGHRAKIEELPFIHPWPFTTPK
KHQEPPLVMYGLKDWKTVQFIELPKESWTNDIQLVKLNWASQIYVGPINVR
GLCKLAKGKALDITVLTETAELEAENREILREPVGYYDPSKDLVAEIQKQCHD
QWTQYIQEPPNLTCKYAKMTATNDVKQITFAVQKIAMESIVWIKGTPKPRPI
QKETWETWTDYQATLPEWENVTPPLVKLWYLEKEPIIGAEITYVYDGAANRETK
IGKAGYVTDGRQKIVSLTETTNQKTELQALQALODSGSEVNIITDSYALGIQAO
PKSESELNDIIEOLIKKRYLVSWPAAHKGICGNEQIDKLVSKIKVLELGLDK
AOEEHEKYSNRAMASEFNPPIVAKEIVASDCOLGAEIHGGVDCSPGIVOLDC
THLEKGIILVAVHVASGYIEAEVPAETQETAYIILKLAGRWPVKVITHDGSNFTG
AAVKAQWAGIQOESGIPYNPQSGVSEMNKELKIIIGQVRDQAEHLKTAVQMAVF
IHNFKPKGIGYYSAGERTIDITATQISKELQKQITKIQNFVYRDSRDPITWKGPA
KVLWKGCAVVIQDNSDIKVVPRPKYKIIIDYQKQAGACVAGKQDED"
281. .1759
/gene="pol"
/codon_start=1
/product="gag polyprotein"
/protein_id="AAD17134.1"
/db_xref="GI:4324870"
/translation="MGAKASILRGEKLDANERIPRPGGKCYMYLKHIVWASRELERF
ALNFKLSEGGKQIVKQALQALQITGIELKSLNVAIYCVHEKIKVKDKKALD
KIEEONKSEKTKQAAAGAKTSQNPVQVQGMVHQAIISPTLNAAVVKVIEKA
FSPEVIMFMTALSEGATPDQNTMTNTVGGHQAAMKLDIINEEAAEWRLHPVHAG
PIACQKREPGSDIAGTSTLOEITMTSNPIIPVGIYKKWIILGLINKIVRMYSF
VSLIDIRQKREPRDYVDFKTLAEQATQDVKNMTDTLLVONANPCKTILRAL
GPATLEEMTACQVGGPGHKARVIAEAMSOATSNILMORSNFKGPKRIKFCNCG
KEGHIARNCRAPKKGCKGEGHQMCKTERQANFFREDLAFPOGKARFEPSEQR
ANSPTRELOVRGDNPRSEAGERTLNFQITLWOPLVSTIKVGGQIKKALDITGA
DDVLEMSLPGKWKPMIGGIGGFTKVRQYDQIIEJCGKKAIGTVIIGPTPPVNIIG
RNMLTQGLTNFPIPIETVPVKLPKMGPKVKOWPITEEIKATITACIEMKEG
KATLGPENYNTVPFAIKKDKKPKLVDEPFIKPTQDFWFOIGIPIHPAGLKKK
KSVTLVDGATFSVLDEDFRKYTAFTIPSINNETPGIRYOYNVLPOCWKSPSIEQ
SSMTKILPEFRANPELVYQMDLYVSDLEIGHRAKIEELPFIHPWPFTTPK
KHQEPPLVMYGLKDWKTVQFIELPKESWTNDIQLVKLNWASQIYVGPINVR
GLCKLAKGKALDITVLTETAELEAENREILREPVGYYDPSKDLVAEIQKQCHD
QWTQYIQEPPNLTCKYAKMTATNDVKQITFAVQKIAMESIVWIKGTPKPRPI
QKETWETWTDYQATLPEWENVTPPLVKLWYLEKEPIIGAEITYVYDGAANRETK
IGKAGYVTDGRQKIVSLTETTNQKTELQALQALODSGSEVNIITDSYALGIQAO
PKSESELNDIIEOLIKKRYLVSWPAAHKGICGNEQIDKLVSKIKVLELGLDK
AOEEHEKYSNRAMASEFNPPIVAKEIVASDCOLGAEIHGGVDCSPGIVOLDC
THLEKGIILVAVHVASGYIEAEVPAETQETAYIILKLAGRWPVKVITHDGSNFTG
AAVKAQWAGIQOESGIPYNPQSGVSEMNKELKIIIGQVRDQAEHLKTAVQMAVF
IHNFKPKGIGYYSAGERTIDITATQISKELQKQITKIQNFVYRDSRDPITWKGPA
KVLWKGCAVVIQDNSDIKVVPRPKYKIIIDYQKQAGACVAGKQDED"

```




[illegible][illegible]

```
Db 1441 ATACAGAACTGATGCTTTAACTTCCTCAGATCACTCTTTGGCAAGACCCCTCG 1500
Qy 1471 aqccagtaa 1479
      | | | | |
Db 1501 TCACAGTAA 1509

RESULT 5
AF287352 4352 bp DNA SYN 01-NOV-2000
DEFINITION Synthetic construct HIV-1-derived gag protein and pol protein
genes, complete cds.
ACCESSION AF287352
VERSION AF287352.1 GI:11066861
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Fuller, M. and Anson, D. S.
TITLE Helper plasmids for production of HIV derived vectors
JOURNAL Unpublished
REFERENCE
AUTHORS Fuller, M. and Anson, D. S.
TITLE Direct Submission
JOURNAL S. Hmitted (12 Jul 2000) Chemical Pathology, Women's and Children's
Hospital, 72 King William Road, North Adelaide, SA 5006, Australia
FEATURES
source
1. 4352
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="derived from human immunodeficiency virus type 1
isolate YU-2"
37. 1539
/notes="codon optimized reading frame"
/codon_start=1
/transl_table=11
/product="gag protein"
/protein_id="AAG28735.1"
/db_xref="GI:11066862"
/translation="MGAPASVLSACGPIDWKPIPIPIPGFFOYPIKHIVWASPELERP
AVSPCLLESEGTGPIIGLQHSITGSEFELSYNTVATLYCVCHQKTEVQDKKALE
KTEENKSSKKKAQAADATGSSQVSONYPIVQNLQGMVHOATSPRTLNAAWKVVE
EKAFSEVIFMFSALSEGATPDQNLNIMNTVGHQAAMQMLKETINEEAADWDLHPV
HAGPIAPQGMREPRGSDIATTTGQIGWMTNNPPIPVGEIYIRWIIIGLNLKLVRM
YSPITLDIPQGPKEPFIYVVPYKTLPAEQASQEVKNMWTETILIVONAPDKTLL
KALGPAATLEMMTACQGVGGPGHKARVLAFAMSVINSAILMMJGKGFHNRKATVKC
FNGKPGHIAKNCRAPKKGCKWCKCGKFGHOMKDCTEROANFLGKIWPSHKGRPGFTQ
SREPTAPSESVRFGEETITPSQKQEPIDKELYPLASLSLSLESGDSPSQ"
<138..4343
/notes="codon optimized reading frame"
/codon_start=1
/transl_table=11
/product="pol protein"
/protein_id="AAG28736.1"
/db_xref="GI:11066863"
/translation="REDLAPQCKAKFSSEOTRANSPRPQVMPDRDNNLSSEAGA
IPWVTSFSPFQITIMQPLVITRTGQIKREALITGCAADTVLEEMNIPIPKPKMIG
GIGFTKQDQDQPIEIGCKAIGIVLGPPIVNLIGRNLIIQICLINFLESIET
VPVKLKPMQPKVKOMPLTEERIKALVEICTEMEKESKISIGPENPVNTPVFAIKK
KDSNPKPLVDFRELKRTDQWELVGLPHFAGLKKKKSVTIVDGTAFYSVLHED
PRKYATFIPIINNETGTROYNLVPCQKGSPIAFQSSMTIILPFKQNPDLVY
QYMDLYVGSLEFLGQHTKIEELRHLIRWGFTTPDKKHKEPPPLMWGYELHFDKW
TVQIVIPFDTSVNIPEKIVSKLWASQIVAGIKVPOICLKHGKALTEVILPIE
EAELEAENRELIKRELVHYDPSKDLIAETOKQSQQWQYQIQEPEPKNLKTQYA
WFGZAHNIVKQI TEAVKPTATESIVTWKTPKPTIPQKPTWFTWMTYQWATWPE
WEFVNTPIVLKWLQLEKEPIIGAEIFYVUGAANRETKLGKAGYVINGKQWVSLTD
TNGKTFLOATVIALQUSQLEVINIVTDSQVAGLIIQAQPDRESEHVSQIIPOLKKE
KVLVWPAHKGIGGNEQDKLVLSAGIKVFLFDGDKAOKRESEKYSNNRWASDPEN
LPFWAKELVASDCKQKLFKPMHGVQDCSPGTLQDCTHLEGVLIIVAHVVASQYIE
AEVIPAGTGTATFLLKLAIRWVLIILHDKNSNFIISALVKAALWKAHLKQFEGIPY
NPQSQVGVSENNKELKLIQVDRQDAELHIAVQMAVFLHNRKKGSGIGYSAQNPVIV
DLIATDIQPKHLQKTIKIQNPRVYVYNDSDPLMKGAKLIDWKGAGVVIQUNSDIKV
```


[illegible][illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2001, 20:59:23 ; Search time 3475.48 Seconds
(without alignments)
4104,284 Million cell updates/sec

Title: us-09-475-704-4

Perfect score: 1509

Sequence: 1 atgaagagagagagagagat.....agagagagagagagagagaa 1509

Scoring table: IDENTITY_NUC

Gapop 10,0 ; Gapext 1,0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456240

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post processing: Minimum Match 9%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: qb_est1:*

2: qb_est2:*

3: qb_est3:*

4: qb_est4:*

5: qb_est5:*

6: qb_est6:*

7: qb_est7:*

8: qb_est8:*

9: qb_est9:*

10: qb_est10:*

11: qb_est11:*

12: qb_est12:*

13: qb_est13:*

14: qb_est14:*

15: qb_est15:*

16: qb_est16:*

17: qb_est17:*

18: qb_est18:*

19: qb_est19:*

20: qb_est20:*

21: qb_est21:*

22: qb_est22:*

23: qb_est23:*

24: qb_est24:*

25: qb_est33:*

26: qb_est34:*

27: qb_est35:*

28: qb_est36:*

29: qb_est37:*

30: qb_est38:*

31: qb_est39:*

32: qb_est40:*

33: em_est10:*

34: em_est11:*

35: em_est12:*

36: em_est13:*

37: em_est14:*

38: em_est15:*

39: em_est16:*

40: em_est17:*

41: em_est18:*

42: em_est19:*

43: em_est20:*

44: em_esthum10:*

45: em_esthum11:*

46: em_esthum12:*

47: em_esthum13:*

48: em_esthum14:*

49: em_esthum15:*

50: em_esthum16:*

51: em_esthum17:*

52: em_esthum18:*

53: em_esthum19:*

54: em_esthum20:*

55: em_esthum21:*

56: em_esthum22:*

57: em_esthum23:*

58: em_esthum24:*

59: em_esthum25:*

60: em_esthum26:*

61: em_esthum27:*

62: em_esthum28:*

63: em_esthum29:*

64: em_esthum30:*

65: em_esthum31:*

66: em_esthum32:*

67: em_esthum33:*

68: em_esthum34:*

69: em_esthum35:*

70: em_esthum36:*

71: em_esthum37:*

72: em_esthum38:*

73: em_esthum39:*

74: em_esthum40:*

75: em_esthum41:*

76: em_esthum42:*

77: em_esthum43:*

78: em_esthum44:*

79: em_esthum45:*

80: em_esthum46:*

81: em_esthum47:*

82: em_esthum48:*

83: em_esthum49:*

84: em_esthum50:*

85: em_esthum51:*

86: em_esthum52:*

87: em_esthum53:*

88: em_esthum54:*

89: em_esthum55:*

90: em_esthum56:*

91: em_esthum57:*

92: em_esthum58:*

93: em_esthum59:*

94: em_esthum60:*

95: em_esthum61:*

96: em_esthum62:*

97: em_esthum63:*

98: em_esthum64:*

99: em_esthum65:*

100: em_esthum66:*

101: em_esthum67:*

102: em_esthum68:*

103: em_esthum69:*

104: em_esthum70:*

105: em_esthum71:*

106: em_esthum72:*

107: em_esthum73:*

108: em_esthum74:*

109: em_esthum75:*

110: em_esthum76:*

111: em_esthum77:*

112: em_esthum78:*

113: em_esthum79:*

114: em_esthum80:*

115: em_esthum81:*

116: em_esthum82:*

117: qb_est48:*
118: qb_est49:*
119: qb_est50:*
120: qb_est51:*
121: qb_est52:*
122: qb_est53:*
123: qb_est54:*
124: qb_est55:*
125: qb_est56:*
126: qb_est57:*
127: qb_est58:*
128: qb_est59:*
129: qb_est60:*
130: qb_est61:*
131: qb_est62:*
132: qb_est63:*
133: qb_est64:*
134: qb_est65:*
135: qb_est66:*
136: qb_est67:*
137: qb_est68:*
138: qb_est69:*
139: qb_est70:*
140: qb_est71:*
141: qb_est72:*
142: qb_est73:*
143: qb_est74:*
144: qb_est75:*
145: qb_est76:*
146: qb_est77:*
147: qb_est78:*
148: qb_est79:*
149: qb_est80:*
150: qb_est81:*
151: qb_est82:*
152: qb_est83:*
153: qb_est84:*
154: qb_est85:*
155: qb_est86:*
156: qb_est87:*
157: qb_est88:*
158: qb_est89:*
159: qb_est90:*
160: qb_est91:*
161: qb_est92:*
162: qb_est93:*
163: qb_est94:*
164: qb_est95:*
165: qb_est96:*
166: qb_est97:*
167: qb_est98:*
168: qb_est99:*
169: qb_est100:*
170: qb_est101:*
171: qb_est102:*
172: qb_est103:*
173: qb_est104:*
174: qb_est105:*
175: qb_est106:*
176: qb_est107:*
177: qb_est108:*
178: qb_est109:*
179: qb_est110:*
180: qb_est111:*
181: qb_est112:*
182: qb_est113:*
183: qb_est114:*
184: qb_est115:*
185: qb_est116:*
186: qb_est117:*
187: qb_est118:*
188: qb_est119:*
189: qb_est120:*

190: qb_est121:*
191: qb_est122:*
192: qb_est123:*
193: qb_est124:*
194: qb_est125:*
195: qb_est126:*
196: qb_est127:*
197: qb_est128:*
198: qb_est129:*
199: qb_est130:*
200: qb_est131:*
201: qb_est132:*
202: qb_est133:*
203: qb_est134:*
204: qb_est135:*
205: qb_est136:*
206: qb_est137:*
207: qb_est138:*
208: qb_est139:*
209: qb_est140:*
210: qb_est141:*
211: qb_est142:*
212: qb_est143:*
213: qb_est144:*
214: qb_est145:*
215: qb_est146:*
216: qb_est147:*
217: qb_est148:*
218: qb_est149:*
219: qb_est150:*
220: qb_est151:*
221: qb_est152:*
222: qb_est153:*
223: qb_est154:*
224: qb_est155:*
225: qb_est156:*
226: qb_est157:*
227: qb_est158:*
228: qb_est159:*
229: qb_est160:*
230: qb_est161:*
231: qb_est162:*
232: qb_est163:*
233: qb_est164:*
234: qb_est165:*
235: qb_est166:*
236: qb_est167:*
237: qb_est168:*
238: qb_est169:*
239: qb_est170:*
240: qb_est171:*
241: qb_est172:*
242: qb_est173:*
243: qb_est174:*
244: qb_est175:*
245: qb_est176:*
246: qb_est177:*
247: qb_est178:*
248: qb_est179:*
249: qb_est180:*
250: qb_est181:*
251: qb_est182:*
252: qb_est183:*
253: qb_est184:*
254: qb_est185:*
255: qb_est186:*
256: qb_est187:*
257: qb_est188:*
258: qb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

em nucleic nucleic search, using sw model

Run on: August 8, 2001, 20:58:59 ; Search time 3475.48 Seconds
(without alignments)
4022.688 Million cell updates/sec

Files: us-09-475-704-3

Perfect score: 1479

Sequence: 1 atqqgcgcacgcacacat.....adgaacacclaaqaadlaa 1479

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 10228115 seqs, 4724426759 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST :

1: qb_est1:*

2: qb_est2:*

3: qb_est3:*

4: qb_est4:*

5: qb_est5:*

6: qb_est6:*

7: qb_est7:*

8: qb_est8:*

9: qb_est9:*

10: qb_est10:*

11: qb_est11:*

12: qb_est12:*

13: qb_est13:*

14: qb_est14:*

15: qb_est15:*

16: qb_est16:*

17: qb_est17:*

18: qb_est18:*

19: qb_est19:*

20: qb_est20:*

21: qb_est21:*

22: qb_est22:*

23: qb_est23:*

24: qb_est24:*

25: qb_est25:*

26: qb_est26:*

27: qb_est27:*

28: qb_est28:*

29: qb_est29:*

30: qb_est30:*

31: qb_est31:*

32: qb_est32:*

33: qb_est33:*

34: qb_est34:*

35: qb_est35:*

36: qb_est36:*

37: qb_est37:*

38: qb_est38:*

39: qb_est39:*

40: qb_est40:*

41: qb_est41:*

42: qb_est42:*

43: qb_est43:*

44: qb_est44:*

45: qb_est45:*

44: em_esthum0:*

45: em_esthum1:*

46: em_esthum2:*

47: em_esthum3:*

48: em_esthum4:*

49: em_esthum5:*

50: em_esthum6:*

51: em_esthum7:*

52: em_esthum8:*

53: em_esthum9:*

54: em_esthum10:*

55: em_esthum11:*

56: em_esthum12:*

57: em_esthum13:*

58: em_esthum14:*

59: em_esthum15:*

60: em_esthum16:*

61: em_esthum17:*

62: em_esthum18:*

63: em_esthum19:*

64: em_esthum20:*

65: em_esthum21:*

66: em_esthum22:*

67: em_esthum23:*

68: em_esthum24:*

69: em_esthum25:*

70: em_esthum26:*

71: em_esthum27:*

72: em_esthum28:*

73: em_esthum29:*

74: em_esthum30:*

75: em_esthum31:*

76: em_esthum32:*

77: em_esthum33:*

78: em_esthum34:*

79: em_esthum35:*

80: em_esthum36:*

81: em_esthum37:*

82: em_esthum38:*

83: em_esthum39:*

84: em_esthum40:*

85: em_esthum41:*

86: em_esthum42:*

87: em_esthum43:*

88: em_esthum44:*

89: em_esthum45:*

90: em_esthum46:*

91: em_esthum47:*

92: em_esthum48:*

93: em_esthum49:*

94: em_esthum50:*

95: em_esthum51:*

96: em_esthum52:*

97: em_esthum53:*

98: em_esthum54:*

99: em_esthum55:*

100: em_esthum56:*

101: em_esthum57:*

102: em_esthum58:*

103: em_esthum59:*

104: em_esthum60:*

105: em_esthum61:*

106: em_esthum62:*

107: em_esthum63:*

108: em_esthum64:*

109: em_esthum65:*

110: em_esthum66:*

111: em_esthum67:*

112: em_esthum68:*

113: em_esthum69:*

114: em_esthum70:*

115: em_esthum71:*

116: em_esthum72:*

117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

190: qb_est121:*
 191: qb_est122:*
 192: qb_est123:*
 193: qb_est124:*
 194: qb_est125:*
 195: qb_est126:*
 196: qb_est127:*
 197: qb_est128:*
 198: qb_est129:*
 199: qb_est130:*
 200: qb_est131:*
 201: qb_est132:*
 202: qb_est133:*
 203: qb_est134:*
 204: qb_est135:*
 205: qb_est136:*
 206: qb_est137:*
 207: qb_est138:*
 208: qb_est139:*
 209: qb_est140:*
 210: qb_est141:*
 211: qb_est142:*
 212: qb_est143:*
 213: qb_est144:*
 214: qb_est145:*
 215: qb_est146:*
 216: qb_est147:*
 217: qb_est148:*
 218: qb_est149:*
 219: qb_est150:*
 220: qb_est151:*
 221: qb_est152:*
 222: qb_est153:*
 223: qb_est154:*
 224: qb_est155:*
 225: qb_est156:*
 226: qb_est157:*
 227: qb_est158:*
 228: qb_est159:*
 229: qb_est160:*
 230: qb_est161:*
 231: qb_est162:*
 232: qb_est163:*
 233: qb_est164:*
 234: qb_est165:*
 235: qb_est166:*
 236: qb_est167:*
 237: qb_est168:*
 238: qb_est169:*
 239: qb_est170:*
 240: qb_est171:*
 241: qb_est172:*
 242: qb_est173:*
 243: qb_est174:*
 244: qb_est175:*
 245: qb_est176:*
 246: qb_est177:*
 247: qb_est178:*
 248: qb_est179:*
 249: qb_est180:*
 250: qb_est181:*
 251: qb_est182:*
 252: qb_est183:*
 253: qb_est184:*
 254: qb_est185:*
 255: qb_est186:*
 256: qb_est187:*
 257: qb_est188:*
 258: qb_est189:*

pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.



QY 712 aqacactgcagagagagatgccttgatgacacagacaccccccgcctgacagagac 771
 Db 1056 AGTACCTTTCAGGAACAGATCGGCTGGATGACCAACACCCACCTAATCGGCTGAGAGAA 1115
 QY 772 atctacaagcgttgatcctctgagcctgaaacaagatcgtgagatgacagcccggtg 831
 Db 1116 ATCTACAAAGCTGGATCATCTCTGGGCTTGAACAAGATGTCGCGATGATAGCCCTAC 1175
 QY 832 agatcctgacatcagcag 891
 Db 1176 AGCAATCGGACATCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1235
 QY 892 ttaagacactgcagcag 951
 Db 1236 TACAAACGCTCCCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1295
 QY 952 ctgctgctgag 1011
 Db 1296 CTGCTGCTGAG 1355
 QY 1012 gcaacactgcag 1071
 Db 1456 GCTACCTTAG 1415
 QY 1072 cggctgctgag 1128
 Db 1416 GCGTCTGAG 1475
 QY 1129 agcaacttcaagggcccccgcgagagagagagagagagagagagagagagagagagag 1188
 Db 1476 GCGAATCTGAG 1535
 QY 1189 atcgcaagaactgcgag 1248
 Db 1536 ACAGCGCCCACTGAG 1595
 QY 1249 caccagatgaag 1308
 Db 1596 CACCAATGAAGAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1655
 QY 1309 cacaag 1368
 Db 1656 CACAG 1698
 QY 1369 ccaag 1428
 Db 1699 -CAACAGCCCAACAG 1757
 QY 1429 gagcacaag 1488
 Db 1758 AAGCAGGAGGAG 1817
 QY 1489 aagcag 1509
 Db 1818 AGCGAGCCCTGCTCACAATAA 1838

RESULT 13
 AF110967
 LOCUS AF110967 5056 bp DNA VRL 02-MAR-2001
 DEFINITION HIV-1 isolate C-96BW05.02 country Botswana, complete genome.
 ACCESSION AF110967
 VERSION AF110967.1 GI:4324723
 KEYWORDS
 SOURCE Human immunodeficiency virus type 1.
 ORGANISM Human immunodeficiency virus type 1
 Viruses; Retroviridae; Retroviridae; Lentivirus, Frimote
 lentivirus group.
 1 (bases 1 to 9056)
 Novitsky, V.A., Montano, M.A., McLane, M.F., Renjifo, B., Vannberg, F.,
 Foley, B.T., Ndung'u, T.P., Rahman, M., Makhele, M.J., Marlink, R. and
 Essex, M.
 TITLE Molecular cloning and phylogenetic analysis of human

immunodeficiency virus type 1 subtype C: a set of 23 full-length clones from Botswana
 J. Virol. 73 (5): 4427-4432 (1999)
 99214383
 10196340
 2 (bases 1 to 9056)
 Novitsky, V.A., Montano, M.A., McLane, M.F., Renjifo, B., Vannberg, F.,
 Foley, B.T., Ndung'u, T.P., Marlink, R. and Essex, M.
 Direct Submission
 Submitted (03-DEC-1998) Immunology and Infectious Diseases, Harvard
 School of Public Health, FXB-310, 651 Huntington Avenue, Boston, MA
 02115, USA
 Location/Qualifiers
 1..9056
 /organism="Human immunodeficiency virus type 1"
 /proviral
 /isolate="C-96BW05.02"
 /db_xref="taxon:11676"
 /country="Botswana"
 /note="subtype: C"
 1..121
 277..4589
 /gene="gag"
 277..4589
 /gene="pol"
 join(277..1560,1560..4589)
 /codon_start=1
 /product="gag-pol polyprotein"
 /protein_id="AADI7072.1"
 /db_xref="GI:424808"
 /translation="MCAPASIIIGFETIKWEKIRLPGGKKHYMLKHLVMSARELGF
 ALNGLLETAEGCKQIMKQLQALQGTGTEELSLYNTVATLYCVHAGIEVDTKEALD
 KIEEUNKSQOKTQAKADGKVSNIPIVNLQGMVLIQALSPRTLANWKVIERKA
 FSEVTPMFTALSEGATPDNLINLNIWGHUAMQMLKDTINEEAEDWRLHPVQAG
 VPAPGMRDPRGSDTAGATSTLQEIAMTNPVPVGDYIKRWIIIGLNKIVRMYSP
 VSLIDIPQSGKEPPFDYVDFPKTLRAEQATQDVKNMTETLLVQNANPDKTILRAL
 GPGATILEMNTACQCGGCGHKARVIAFAMSCANSVNIIMOKSNFKGPRRVKFCNG
 KEGHIAKNCAPKKCKWCKGKHUKCTERQANFRENLAPOKAREFPVQIR
 ASSPNSNSTPRELQPDNDPPSPAGAFGGCTAGCTLNCQITLWQRLVSIKVGGO
 IKAALDGTADNITVLEINLPKWKPKMTIGGIGGFKVRQDYQIVFICGKAGITVL
 VPTPVNIIGRNMTIQLGCTLNPIPIETVFKLPGMDGPKQWQVTLTEKIKALT
 EICEEMEKGGKTKIGPENPYNTPVFAIKKDKTKRWLVDFRELKRLQIEVWVQ
 HTPACGLKKKSVTVLDMDGDAVSPVLDGPKTYATFIPSTNNETPGIRYVRLPQ
 GNKSPALFQSSMTKILFELRNQPELVLYMDULYVGSLEIGSPPAQIEELRHLA
 IAWGFTTPQKKQKPEHPMGLYELHPDKTVPQIQLPKDKSWIVNDIKLVGKINWA
 SUIYGIKVKQLCKLLRGAALTDVVPILTEAELEAENREILKEPVHGVYDPSKL
 IATIKQGHQDWTYQIYQEPFKNLTKGTAKMRTAHTNDVKQLTEAVOKIAQESIVIH
 GKTPEKPLIQETWETWTWYQATWIPENEFVNTPLVLWYOLEKEPIPGVEIFY
 VDAANRETGLGAGVYTDGRQKIVSLTETTNKTELOATOLALQDSQSEVNIYDS
 QYALGILQAPUKSESELVNLIEULQKEMWVLSWIFAIHKGIGNEQWDLKVSQIR
 KMLPDLQIDKAPQEPHEKYHNNRMADENLPIVAKETVASCDCOLKGFALHGOVD
 CSPGILQIDCTHLEGGKVLVAVHVASGYMEAEVIPAETGPTAYFILKAGRWPRVI
 HTDNGTNTSAVAAKCAWAGIQQEPGIPYNPQSGGVESNMKELKKIIGQVRQAEPH
 LKTAQMAVFIHNFPRKGGIGGYSAGERIIDIIATDIQIKELQKQIILKIQFRVYKRD
 SRDPIWKGPAKLLWKGGAVTQDNSDIKVPVRKAKIKDYGKMGAGACVAGGQEN
 N"
 277..1785
 /gene="pol"
 /codon_start=1
 /product="gag polyprotein"
 /protein_id="AADI7071.1"
 /db_xref="GI:4324807"
 /translation="MGARASIIIGKEKDKWEKIRLPGGKKHYMLKHLVMSARELGF
 ALNGLLETAEGCKQIMKQLQALQGTGTEELSLYNTVATLYCVHAGIEVDTKEALD
 KIEEUNKSQOKTQAKADGKVSNIPIVNLQGMVLIQALSPRTLANWKVIERKA
 FSEVTPMFTALSEGATPDNLINLNIWGHUAMQMLKDTINEEAEDWRLHPVQAG
 VPAPGMRDPRGSDTAGATSTLQEIAMTNPVPVGDYIKRWIIIGLNKIVRMYSP
 VSLIDIPQSGKEPPFDYVDFPKTLRAEQATQDVKNMTETLLVQNANPDKTILRAL
 GPGATILEMNTACQCGGCGHKARVIAEAMSAQSNVNIIMOKSNFKGPRRVKFCNG
 KEGHIAKNCAPKKCKWCKGKHUKCTERQANFRENLAPOKAREFPVQIR
 PAAPTVPAPALSPFEETTPAPKPEKDPREPPELTAALSLFSGSPLSQ"
 4534..5112
 /gene="vif"


```

DE 1717 AGCGAAGCTACAGGACACCTTAACAGCGGCTCACATCCTCTTGAGAGGCGCCCTTG 1776
QY 1501 agcagaa 1509
    || |||
DB 1777 TCCTCAATAA 1785

RESULT 15
AF286234          9034 bp      DNA          VRL          10-APR-2001
LOCUS            HIV-1 strain 98T2013 from Tanzania, complete genome.
DEFINITION
ACCESSION        AF286234
VERSION          AF286234.1 GI:13569317
KEYWORDS
SOURCE            Human immunodeficiency virus type 1.
ORGANISM          Viruses: Retroid viruses; Retroviridae; Lentivirus; Primate
                  lentivirus group.
REFERENCE
  1 (bases 1 to 9034)
  Robinson,C.M., Li,Y., Trask,S.A., Chen,Y., Decker,J.,
  Robertson,D.J., Kallish,M.L., Shaw,G.M., Allen,S., Hahn,B.H. and
  Gao,F.
  Near full-length clones and reference sequences for subtype C
  isolates of HIV type 1 from three different continents
  AIDS Res. Hum. Retroviruses 17 (2), 161-168 (2001)
  21094715
  2 (bases 1 to 9034)
  Robinson,C.M., Li,Y., Trask,S.A., Chen,Y., Decker,J.,
  Robertson,D.L., Allen,S., Shaw,G.M., Hahn,B.H. and Gao,F.
  Direct Submission
  Submitted (11-JUL-2000) Medicine, University of Alabama at
  Birmingham, 701 19th Street, South, UHRB 641, Birmingham, AL 35294,
  USA
FEATURES
  source
    1..9034
    /organism="Human immunodeficiency virus type 1"
    /proviral
    /strain="98T2013"
    /db_xref="taxon:11676"
    /country="Tanzania"
    /note="isolated in 1998"
    136..1635
    /gene="gag"
    136..1635
    /gene="gag"
    /codon_start=1
    /product="gag protein"
    /protein_id="AAK31053.1"
    /db_xref="GI:13569318"
    /translation="MGARASILKGGKLDKWEKIRLPDGGKKHYMIKHLVWASRELERF
    ALNPSLETADGCKQIRQLHPALQGTIEELRSLENTVATLYCVHAEIDVDRDTEALD
    KIEEQNKSOQTKQAALVKISQNPVIVNLRGQMVHQAIISPTLNAMVKVIEEKA
    FSEVIMPTALSEGATPODLNMLNTVGHQAAMQMLKDTINEEAAEWDRLHPVHA
    PIAPGMRKPKGSDIACTTSTLQGTIAWMTGNPPVPGVDYKRWIIIGLNIVRMYSI
    VSILDIQGPKEPRDYDRFPKTPAEQATODVKNWMTDLIVONANPCKTILRAL
    GPAATLEEMTACGGVGGSHKARVLAEAMSQTTSTNIMORSNFKPKRTVKCFNCQ
    KEGHIANCPAPKSKCKWKGRHGMKGRTEFANFLKRWPSHKGRPGQFLQGRPE
    PTAPAPPTAPPAESFRFEGTTPAPKQEPKDEPLTSLKSLFGNDPLFQ"
    <1419..4439
    /gene="pol"
    <1419..4464
    /gene="pol"
    /codon_start=1
    /product="truncated pol polyprotein"
    /protein_id="AAK31054.1"
    /db_xref="GI:13569319"
    /translation="FTFENIAPPGGEARKFPSPGAPANSPTSPANSPTSRRLVGRDIN
    PRSETGARGCTNLQALWQRPPLYSIKVGGVKEALLDTCADDTVLFEINLPCKWK
    PKMTGGGKIKKVRQVDQIPDIFICGKALGTVLGTPVNIIGRNMLTGLGCTLNFP
    SPIETGVKILPGMDGPKVQWPLTEEKIKALTAICEEMEKEGKTIKIGPENPTYTV
    FAIKKDSQTKRWLVDFRELNRKTDQFWEVQIIPHAPGLKKKRSVILVLEQDFTSV
    PLDGGPKRYTAFTIPSTNNETPGIRYOYNVLPQWKGSPAIFQSSMTKILEPFIQNP
    GIVTIYQMDLDLYVGSDELEIGOHAKRVRELRLKWKGTTPDKKHQKEPFPFLMGVEL

```

gene
CDS

gene
CDS

gene
CDS

gene
CDS

gene
CDS

gene
CDS

```

HPDKWTVPQIQLPKESWTVNDIOKLVGKLNWASQIHSGTIKHQICKLILRGAKALTDI
VPLIKAEALLEENREITIKPEVHVGVYDPSKULIAEIQKQGGQWTVYQIQEFPNCK
TGKVKMRSANTNDVKOLTEAIQKIAMESIVIWGKTPEKFLPIQKETWETWITDYOQA
TWIPEFEVANTEPLVKWGLEKETAGATYVVDGAANPESIKGKAGYITDPGPKI
ITLAETTNQAKELQALQALQDSCSKVNIATDSYALGIIQADPKSELSYQIIEQ
LIGKERVTLISWVPAHKGIGGNEQVDKLVSSGIRKVLFDGIDKAOEHEKYSNWFAM
ASEFNLPPIVAKETIVASCDCQKLGKGAIHGGVDCSPCIWOLUICHTLEGKILVAVHA
SGYTEAEVIPAFTQETA"
4384..4962
/gene="vif"
4384..4962
/gene="vif"
/codon_start=1
/product="vif protein"
/protein_id="AAK31055.1"
/db_xref="GI:13569320"
/translation="MENRQVLIWQVDKMKIKTWSLVKYHMYISPRASGWFYPIHY
KSRHPKVSSEVHIPIGGAIVITTYMGLQGTGEGWHLGHGVSIEWRLKRYSTOVDSI.
ADQLIHMYFDICFADSAIRKALGSHIVIPRCDYQAGINKVCSLOYLALITALIKPKPK
PPLPSVPLKLVPRNNPQKTRGRGNHTMSGH"
4902..5192
/gene="vpr"
4902..5192
/gene="vpr"
/codon_start=1
/product="vpr protein"
/protein_id="AAK31056.1"
/db_xref="GI:13569321"
/translation="MEQVPEQGPQREFYNEWALEILEDLKQEAIVHFPRPWLHDLQ
HIVVNYGDIWIGVEALHKLQULLFIIHKIGCUHSRIGILKQRARRAGASRS"
5173..7823
/gene="tat"
join(5173..5387,7733..7823)
/gene="tat"
/codon_start=1
/product="tat protein"
/protein_id="AAK31057.1"
/db_xref="GI:13569322"
/translation="MEPVDPKLEPNHPSOPKTECNKCYKHCYSYHCLVCFOTKGLG
ISYGPKKPPQPPSTPPSSDHQNLISKPLSTQDQGTGPEQKKVESKTKTDPPD"
5312..7980
/gene="rev"
join(5312..5387,7733..7980)
/gene="rev"
/codon_start=1
/product="rev protein"
/protein_id="AAK31058.1"
/db_xref="GI:13569323"
/translation="MAGEFSGTSDEALLRAVRIIKILYQSNPYPEPRGTPQACKNPPRR
WPAAPQIHHSISFRILSSGLGRSAEPVPLQLPPIERLYIDCSGRTSGTQSQGTTE
GVGSP"
5401..5661
/gene="vpu"
5401..5661
/gene="vpu"
/codon_start=1
/product="vpu protein"
/protein_id="AAK31059.1"
/db_xref="GI:13569324"
/translation="MINLQARDIYRLGVALIVALSIAIVVMITAYLEVPKLVPQPKI
DWLIKPIREADSGNSEGVTEELATWMDMGLKLLNNDL"
5579..8170
/gene="env"
5579..8170
/gene="env"
/codon_start=1
/product="truncated envelope glycoprotein"
/protein_id="AAK31060.1"
/db_xref="GI:13569325"
/translation="MEVPQIPRNQPPWWTWIGLGFMTMTCNMVGNLWTVYGVYVW
KEAKTTLPCASDAKQYKRVHNVWATHACVPTDPNPOEIVLDNVTENFMKNDKVEQ
MHEDIISIMQSLKPCVKLPLCVTLNCAVATATNGSDIMNTISNMGKMSFN
ITTELKDKKOKKVSALFYKLDIVPLNKNNSFGEYRLINCNTSTITQACPKVSEDDPI

```

```

1414 ---agcaagaccgcgacacctaccagcctgaagagccctgttoggaaacgccccctcg 1470
      ||||| | | ||||||||| | ||||||||| ||||| ||||||||| |||||
Db    4596 atcgaagaaggactgtaccctcctgaccagctgcgcgcctgttcggcaacgcccagc 4655

QY    1471 agccagtaa 1479
      |||||||
Db    4656 agccagtaa 4664

RESULT          9
AAAT0470
ID   AAA70470 standard; DNA; 4766 BP.
AC   XX
CC   AC
XX   AAA70470;
DT   28-Nov-2006 (first entry)
XX
DE   HIV bicistronic construct gp160.modUS4.Gag.modSF2.
XX
KW   HIV-1; AIDS; Gag; vaccine; expression cassette; Env; ss.
QS   Chimeric - Cytomegalovirus.
GS   Chimeric - Human immunodeficiency virus type 1.
XX
PN   WO200003902-A2.
XX
PP   06-JUL-2000.
XX
PF   30-DEC-1999; 99WO-US31245.
XX
PR   31-DEC-1998; 98US-0114495.
PK   01-DEC-1999; 99US-0168471.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Barnett S., Zur Megede J., Srivastava I., Lian Y., Hartog K., Liu H.;
PJ   Greer C., Selby M., Walker C.;
XX
DR   WPI: 2000-45400/39.
XX
PT   Expression cassettes encoding the human immunodeficiency virus (HIV)
PF   Gag-containing polypeptide useful for vaccinating against HIV
PI   infections and acquired immunodeficiency syndrome (AIDS) -
XX
PS   Claim 27; Fig 61; 39pp; English.
XX
AB   The present invention relates to synthetic HIV Gag and Env expression
CC   cassettes. The Gag protein of HIV is needed for the assembly of
CC   virus-like particles. In addition, the Gag protein is involved in many
CC   stages of the HIV life cycle, including assembly, virion maturation at
CC   particle release and early post-entry steps in viral replication. The
CC   expression cassettes may be used for the recombinant expression of HIV
CC   gag- and env-polypeptides which may then be used to vaccinate against
CC   infection and acquired immunodeficiency syndrome (AIDS). The present
CC   sequence is a synthetic construct constructed for the generation of the
CC   expression cassettes of the present invention. This construct is
CC   bicistronic in that the coding sequences for Env and Gag are under the
CC   control of a single CMV promoter and between the two coding sequences
CC   Internal Ribosome Entry Site (IRES) is present.
XX
SQ   Sequence 4766 bp; 1051 A, 1638 C, 1409 G; 668 T; 0 other;

Query Match             81.0%; Score 1197.6; DB 21; Length 4766.
Best local similarity    89.3%; Pred. No. 3e-150;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps

QY    1 atggggccccgcgcacatctctgcggcgcgaagtgcagccctggagcgcattccgc 60
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    3233 atgggcgccgcgcagctgctgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3299

QY    61 ctggccccgcgcgcgaagtgttatcatgatgaacacacctggtgtgggccagccgcag 120
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible]

129 352
 121 180
 135 3412
 181 240
 1413 3472
 241 300
 1473 3532
 301 360
 353 3592
 361 408
 3593 3652
 409 468
 4653 3712
 469 528
 4713 3772
 529 588
 4773 3842
 589 648
 4833 3892
 649 768
 4893 3952
 709 768
 4953 4012
 769 828
 4013 4072
 829 888
 4073 4132
 889 948
 4133 4192
 949 1608
 4193 4252
 1009 1068
 4253 4312
 1069 1122
 4313 4372
 1123 1182
 4373 4432

1183 1242
 4433 4492
 1243 1302
 4493 4552
 1303 1362
 4553 4612
 1363 1413
 4613 4672
 1414 1470
 4673 4742
 1471 1479
 4733 4741

RESULT 10

AAA70415

ID AAA70415 standard; DNA; 2031 BP.

XX AAA70415;

XX 28 NOV 2000 (first entry)

XX Synthetic HIV Gag/NCV core fusion coding sequence.

XX HIV 1; AIDS; Gag, vaccine; expression cassette; ss.

XX Chimeric - Human immunodeficiency virus type 1.

XX Chimeric - Hepatitis C virus.

XX WC200039302-A2.

XX 06-JUL-2000

XX 30-DEC-1999; 99W0-US31245.

XX 31-DEC-1998; 98US-0114495.

XX 01-DEC-1999; 99US-0168471.

XX (CHIK) CHIKON CORP.

XX Barnett S, Zur Meysede J, Srivastava I, Lian Y, Hartog K, Liu H;

XX Greer C, Selby M, Walker C;

XX WP1; 2000-452400/39.

XX Expression cassettes encoding the human immunodeficiency virus (HIV)

XX Gag-containing polypeptide useful for vaccinating against HIV

XX infections and acquired immunodeficiency syndrome (AIDS) -

XX Example 1; Pages 341-342; 391pp; English.

The present sequence is a HIV Gag/Hepatitis C virus (HCV) core fusion coding sequence. The Gag protein of HIV is needed for the assembly of virus-like particles. In addition, the Gag protein is involved in many stages of the HIV life cycle, including assembly, virion maturation, particle release and early post-entry steps in viral replication. The present invention relates to synthetic HIV Gag expression cassettes. The present sequence was cloned and used to generate the expression cassettes of the present invention. The expression cassettes may be used for the recombinant expression of HIV Gag-polypeptides which may then be used to vaccinate against HIV infection and acquired immunodeficiency syndrome.

[illegible]

28 NOV-2000 (first entry)

28 NOV-2000 (first entry)

Thu Aug 9 11:08:08 2001

```

|||||
Db 841 tccatctggacatccgcagagcccccagagagccctccgcactacgtgacacgttc 900
QY 895 ttcaagacctgcgcgcagcagagcaccacagaggtgaagaactgatatccacgacacc 954
Db 901 tacaagacctgcgcgcagcagcctccacagaggtaaagaactgatatccacgagacc 960
QY 955 ctgctggtgcagaaccccaacccacactcgaagaccatcctgcgcgtctcggcccccgc 1014
Db 961 ctgctggtgcagaaccccaacccacactcgaagaccatcctgcgcgtctcggcccccgc 1020
QY 1015 qccagcctgagagagatgatacgcctgcacagggcgtgagcgcgccacgacacagacc 1074
Db 1021 gccacctgagagagatgatacgcctgcacagggcgtgagcgcgccacgacacagacc 1080
QY 1075 cgcgtgcctgcagagcgtgagcagggcccaaca-----ccagcgtatgatgcagaag 1128
Db 1081 cgcgtgcctgcagagcgtgagcagggcccaaca-----ccagcgtatgatgcagaag 1140
QY 1129 agcaacttcagagggcccccgcgcacatcgtaactgcacagcagagagggccac 1188
Db 1141 ggcaccttcgcagacccgcagagacccgtgaggtgcttcaactgcgcagagagggccac 1200
QY 1189 atgcgcgcactgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1248
Db 1201 atgcgcgcactgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1260
QY 1249 caccagatgagagactgcacgcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1308
Db 1261 caccagatgagagactgcacgcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1320
QY 1309 cacaagggccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1368
Db 1321 cacaagggccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1380
QY 1369 gagaqcttcgccttcgagga-----gaccaccccccgcagagcagagag-----agc 1416
Db 1381 gagaqcttcaggtttgggagagagacacacacacacacacacacacacacacacacacac 1440
QY 1417 aaggaccgcagacccctgaccagcctgaaagagcgtgttcgacacacacacacacacacacac 1476
Db 1441 aaggacgtatcctttagcttccctcagatcactctttggcagcgcacccctcgtcacaa 1500
QY 1477 taa 1479
Db 1501 taa 1503

```

Search completed: August 8, 2001, 18:25:56
Job time: 12660 sec